

Access DB# 39046

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched.
 Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or
 utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if
 known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

STAFF USE ONLY

Searcher: AB5501
 Searcher Phone #: 308-4501
 Searcher Location: Biotek Lab
 Date Searcher Picked Up: 3/30/01
 Date Completed: 4/4/01
 Searcher Prep & Review Time: _____
 Clerical Prep Time: 3 m
 Online Time: 2 m

Type of Search

NA Sequence (#) 1
 AA Sequence (#) _____
 Structure (#) _____
 Bibliographic _____
 Litigation _____
 Fulltext _____
 Patent Family _____
 Other _____

Vendors and cost where applicable

STN _____
 Dialog _____
 Questel/Orbit _____
 Dr.Link _____
 Lexis/Nexis _____
 Sequence Systems AB5501
 WWW/Internet _____
 Other (specify) _____

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2001, 21:53:18 ; Search time 100.87 Seconds
(without alignments)
5139.432 Million cell updates/sec

Title: US-09-582-779A-1

Perfect score: 1380

Sequence: 1 ctcgagcaactcattggaag.....aataatgacgttgatc 1380

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_36:*

- 1: /SID56/gcgdata/geneseq/geneq/NA1980.DAT:*
- 2: /SID56/gcgdata/geneseq/geneq/NA1981.DAT:*
- 3: /SID56/gcgdata/geneseq/geneq/NA1982.DAT:*
- 4: /SID56/gcgdata/geneseq/geneq/NA1983.DAT:*
- 5: /SID56/gcgdata/geneseq/geneq/NA1984.DAT:*
- 6: /SID56/gcgdata/geneseq/geneq/NA1985.DAT:*
- 7: /SID56/gcgdata/geneseq/geneq/NA1986.DAT:*
- 8: /SID56/gcgdata/geneseq/geneq/NA1987.DAT:*
- 9: /SID56/gcgdata/geneseq/geneq/NA1988.DAT:*
- 10: /SID56/gcgdata/geneseq/geneq/NA1989.DAT:*
- 11: /SID56/gcgdata/geneseq/geneq/NA1990.DAT:*
- 12: /SID56/gcgdata/geneseq/geneq/NA1991.DAT:*
- 13: /SID56/gcgdata/geneseq/geneq/NA1992.DAT:*
- 14: /SID56/gcgdata/geneseq/geneq/NA1993.DAT:*
- 15: /SID56/gcgdata/geneseq/geneq/NA1994.DAT:*
- 16: /SID56/gcgdata/geneseq/geneq/NA1995.DAT:*
- 17: /SID56/gcgdata/geneseq/geneq/NA1996.DAT:*
- 18: /SID56/gcgdata/geneseq/geneq/NA1997.DAT:*
- 19: /SID56/gcgdata/geneseq/geneq/NA1998.DAT:*
- 20: /SID56/gcgdata/geneseq/geneq/NA1999.DAT:*
- 21: /SID56/gcgdata/geneq/geneq/NA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380	100.0	1380	20	X88501
2	370.4	26.8	2330	17	T08603
3	370.4	26.8	2330	19	V11505
4	369.6	26.8	1469	15	O66476
5	357	25.9	5181	20	X85877
6	357	25.9	7063	20	X85878
7	356	25.8	1101	20	V83956
8	356	25.8	4102	19	V64257
9	356	25.8	4164	16	T04575
10	356	25.8	4164	19	V22271
11	356	25.8	4164	20	V69740
12	356	25.8	4292	20	X60602

13	356	25.8	4933	16	T04569	Plasmid pAT-2 sequ
14	356	25.8	4933	20	V22272	pAT-2 (pSD543). S
15	356	25.8	4933	19	V69741	Nucleotide sequenc
16	356	25.8	5848	19	V50471	Yeast expression v
17	356	25.8	5848	21	Z86938	Recombinant C reac
18	356	25.8	7102	20	X31836	Nucleotide sequenc
19	356	25.8	7333	20	X31837	Nucleotide sequenc
20	356	25.8	7338	19	V34061	Plasmid pSF248 enc
21	356	25.8	7633	20	Z27543	Yeast expression p
22	356	25.8	8140	19	V23013	Nucleic acid sequ
23	355	25.7	804	20	V36432	URA3. Saccharomy
24	355	25.7	804	20	Z07052	Saccharomyces cere
25	355	25.7	804	20	X60601	S. cerevisiae URA3
26	352.8	25.6	1115	19	O71368	URA3-J Allele (sel
27	351.2	25.4	1179	15	V19462	C. utilis URA3 gen
28	349.6	25.3	6978	10	N90096	Sequence of plasmid
29	349	25.3	1220	11	O05864	Penicillium chryso
30	336.4	24.4	906	11	O05865	Vector pPR70 conta
31	289.4	21.0	2688	11	O29938	Construct EC21. (Co
32	273.2	19.8	6640	21	A26729	A. nidulans PCR de
33	270	19.6	7175	21	A26728	Human colon cancer
34	116	8.4	1403	11	O03351	Sequence of the Ta
35	115.2	8.3	7432	19	V61459	Deleya halophila m
36	115.2	8.3	8528	14	O46249	Rhizobium species
37	114.2	8.3	2229	20	V72393	Rhizobium species
38	95.6	6.9	504	21	A16205	Rice beta-glucanase
39	75.8	5.5	444	8	N70493	Rice Gm7 CDS. Or
40	39.8	2.9	753	20	V81457	Rice beta-glucanase
41	39.8	2.9	534720	19	V30458	Rice beta-glucanase
42	39.8	2.9	536165	19	V30459	Rice beta-glucanase
43	38.6	2.8	1008	20	X33569	Rice beta-glucanase
44	38.6	2.8	1008	20	X33569	Rice beta-glucanase
45	38.6	2.8	2810	20	X33535	Rice beta-glucanase

ALIGNMENTS

RESULT 1	
ID X88501	
XX X88501 standard; DNA; 1380 bp.	
AC X88501;	
XX	
DT 06-OCT-1999 (first entry)	
XX	
DE	
XX	
KW Orotidine-5-phosphate decarboxylase; selectable marker; biosynthesis;	
KW Metcchnokowiacae; regulatory activity; chemical biosynthesis;	
KW secondary metabolite biosynthesis; protein biosynthesis; flavouring;	
KW growth factor; perfume; enzyme biosynthesis; riboflavin; vitamin B2;	
KW nutritional supplement; food dye; ss.	
XX	
OS Ashbya gossypil.	
XX	
FH Key	Location/Qualifiers
FT 5'UTR	1..209
FT	/*tag- a
FT	210..1013
FT CDS	/*tag- b
FT	/product- "orotidine-5-phosphate decarboxylase"
FT	1013..1380
FT	/*tag- c
XX	
PN DEL9801120-A1.	
XX	
PD 22-JUL-1999.	
XX	
PF 15-JAN-1998;	98DE-1001120.
XX	
PR 15-JAN-1998;	98DE-1001120.
XX	

PA (BADI) BASF AG.
 XX
 PI Pompejus M. Revuelta Doval JL. Santos Garcia MA:
 XX WPI, 1999-406041/35.
 DR P-PSDB: Y25930.
 XX
 PT New orotidine-5'-phosphate decarboxylase gene useful as selectable
 marker for microbial transformation
 PS
 XX Claim 1: Page 9-10; 14pp: German.
 XX
 CC This invention describes a novel *Ashbya gossypii* orotidine-5'-phosphate
 CC decarboxylase gene (r). (I) is a selectable marker for use during the
 CC introduction of genes into microorganisms, especially those of the family
 CC Metazoa, specifically genera *Exomethecia*, *Ashbya* or
 CC *Nematostella*. The introduced genes may have regulatory activity or may be
 CC involved in the biosynthesis of chemicals, secondary metabolites,
 CC proteins, flavourings, growth factors, perfumes or enzymes. Most
 CC particularly they are involved in the biosynthesis of riboflavin or
 CC Vitamin B2, which is used as nutritional supplement in humans and animals
 CC and as a food dye. (I) is a marker that is readily selectable,
 CC transforms with higher efficiency than thr, leu or kan, and is simply
 CC counter-selected. When used in *Ashbya gossypii* it does not involve
 CC introduction of any foreign DNA. This sequence encodes the A. gossypii
 CC orotidine-5'-phosphate decarboxylase protein.
 CC
 XX
 SO Sequence 1380 BP: 334 A: 354 C: 375 G: 317 T: 0 other:

Query Match 100.0%; Score 1380; DB 20; Length 1380;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctgaggaactcatctggaagcccttcgcaagaccctatctctgcttcaagtctcta 60
 DB 1 ctggaactcatctggaagcccttcgcaagaccctatctctgcttcaagtctcta 60
 QY 61 ctctcgtatgctgctcactacagaaattcttctctatgctgcaagaaagacatca 120
 DB 61 ctctcgtatgctgctcactacagaaattcttctctatgctgcaagaaagacatca 120
 QY 121 catacattctgagtgtagtgcacacacacacacacacacacacacacacacacac 180
 DB 121 catacattctgagtgtagtgcacacacacacacacacacacacacacacacacac 180
 QY 181 taagggtgctacgagccttagcaccatgctcacaatgctcacaagaaagggcca 240
 DB 181 taagggtgctacgagccttagcaccatgctcacaatgctcacaagaaagggcca 240
 QY 241 aggcac 300
 DB 241 aggcac 300
 QY 301 atctctgagccttcctctgtagtgcagacgtctagaagctctctgagcagacacac 360
 DB 301 atctctgagccttcctctgtagtgcagacgtctagaagctctctgagcagacacac 360
 QY 361 tgggacgcgacacattgctgctgtagaacaacatgctgacatactgacgagcttcgacatcg 420
 DB 361 tgggacgcgacacattgctgctgtagaacaacatgctgacatactgacgagcttcgacatcg 420
 QY 421 agagcagcagtcagacgctgagcagacgtctgagccttagaacaacatctcatgcttcgagag 480
 DB 421 agagcagcagtcagacgctgagcagacgtctgagccttagaacaacatctcatgcttcgagag 480
 QY 481 accgcaagtcgctgtagacgtgacacacacacacacacacacacacacacacacac 540
 DB 481 accgcaagtcgctgtagacgtgacacacacacacacacacacacacacacacacac 540
 QY 541 gtaacgagagtgaggcagatattacacacacacacacacacacacacacacacacac 600
 DB 541 gtaacgagagtgaggcagatattacacacacacacacacacacacacacacacacac 600

QY 601 ccgggctcgaagagagctcgcgaacatggtccctacacagaaaccacaggggtctgctgatgctg 660
 DB 601 ccgggctcgaagagagctcgcgaacatggtccctacacagaaaccacaggggtctgctgatgctg 660
 QY 661 cagagctctctctcagagagctctcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 720
 DB 661 cagagctctctctcagagagctctcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 720
 QY 721 tggcgaagctcgaagagagctcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 780
 DB 721 tggcgaagctcgaagagagctcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 780
 QY 781 gtgcgaagagctctgagcagacatctgacacacacacacacacacacacacacacacacacac 840
 DB 781 gtgcgaagagctctgagcagacatctgacacacacacacacacacacacacacacacacacac 840
 QY 841 accgcttgggacacagacatctgac 900
 DB 841 accgcttgggacacagacatctgac 900
 QY 901 tcatgtgtgagagagagctctcttgcacacacacacacacacacacacacacacacacacac 960
 DB 901 tcatgtgtgagagagagctctcttgcacacacacacacacacacacacacacacacacacac 960
 QY 961 accgcaagagcgcgtctggaagagctcttgcacacacacacacacacacacacacacacacac 1020
 DB 961 accgcaagagcgcgtctggaagagctcttgcacacacacacacacacacacacacacacacac 1020
 QY 1021 ctggcgccacagatataagcgcgacatccacacacacacacacacacacacacacacacacac 1080
 DB 1021 ctggcgccacagatataagcgcgacatccacacacacacacacacacacacacacacacacac 1080
 QY 1081 taattatagccctatctgac 1140
 DB 1081 taattatagccctatctgac 1140
 QY 1141 gcaacaagttagatattctgtagtgcagtagacatgacacacacacacacacacacacacacac 1200
 DB 1141 gcaacaagttagatattctgtagtgcagtagacatgacacacacacacacacacacacacacac 1200
 QY 1201 aagccttcgc 1260
 DB 1201 aagccttcgc 1260
 QY 1261 atcttcaatttggccttaactgagcacaatactgacacacacacacacacacacacacacacac 1320
 DB 1261 atcttcaatttggccttaactgagcacaatactgacacacacacacacacacacacacacacac 1320
 QY 1321 agcttggaactaagtgac 1380
 DB 1321 agcttggaactaagtgac 1380

RESULT 2
 T08603
 ID T08603 standard; DNA; 2330 BP.
 XX
 AC T08603;
 XX
 DT 03-DEC-1996 (first entry)
 XX
 DE UR3 gene.
 XX
 KW Ribosomal protein; expression vector; heterogene; yeast; GAP; PGK;
 KW pMA; phosphoglyceric acid kinase; protoplasmic membrane proton ATPase;
 KW glyceraldehyde-3-phosphoric acid dehydrogenase; resistance marker;
 KW hygromycin B phosphotransferase; aminoglycoside-3'-phosphotransferase;
 KW ss.
 XX
 OS Candida utilis.
 XX
 FH Key Location/Qualifiers

FT CDS 1259..2062
 FT /*tag= a
 FT /product= URA3 gene product.
 PN MO9532289-A1.
 PD 30-NOV-1995.
 XX
 XX
 XX 25-MAY-1995; 95MO-JP01005.
 XX
 XX 28-APR-1995; 95JP-0129287.
 PR 25-MAY-1994; 94JP-0135015.
 PR 26-OCT-1994; 94JP-0285823.
 XX
 XX (KIRI) KIRIN BEER KK.
 PA (KIRI) KIRIN BREWERY KK.
 XX
 PI Kajiwara S, Kondo K, Miura N:
 DR WPI; 1996-020584/02.
 DR P-PSDB; R77659.
 XX
 XX Ribosomal protein L41 gene, promoter and terminator sequences,
 PT isolated from Candida utilis - used in vector for expression of
 PT hetero:gene(s) in yeast
 XX
 PS Claim 12; Page 153-156; 252pp; Japanese.
 XX
 CC The L41 ribosomal structural protein gene (708602) can be used in the
 CC construction of expression vectors designed specifically to express
 CC heterologues in yeast. The vectors also comprise: a promoter and
 CC terminator sequence selected from phosphoglyceric acid kinase (Pek),
 CC glyceraldehyde-3-phosphoric acid dehydrogenase (GAP) and
 CC protoplasmic membrane protein ATPase (PMA) genes isolated from
 CC Candida utilis; a drug resistance marker selected from
 CC aminoglycoside-3'-phosphotransferase or hygromycin B
 CC phosphotransferase; and the heterogene to be expressed. L41 confers
 CC cyclohexamine resistance on the yeast. The expression vector may
 CC also contain C. utilis rRNA gene cluster sequences,
 CC orchidin-5'-phosphate decarboxylase, or URA3 gene sequences.
 XX
 SQ Sequence 2330 BP; 669 A; 498 C; 613 G; 550 T; 0 other;

Query Match 26.8%; Score 370.4; DB 17; Length 2330;
 Best Local Similarity 67.5%; Pred. No. 1.9e-96;
 Matches 521; Conservative 0; Mismatches 251; Indels 0; Gaps 0;

QY 214 caaggaatcttcgcaagaagcgcaagcacacaatctgcagcttcgcaagaagcttc 273
 DB 1266 ccacgttatcgtacacagaagagcgatcgacaccccttcgcaactgtcctaagcgcttgc 1325
 QY 274 tggcatgtatgcacgaagaacacaaatctcgcgcttccttcgtatgtcgaagcttca 333
 DB 1326 ttccgcttatcgtgagtcacagaagaacgaactcgtgtcgcagtgctgtcgtacacag 1385
 QY 334 gaaagctctcggagctcagacacacgctcggaccgacacattgtcgtcgtgaagacacat 393
 DB 1386 aagagttgcctcaagctcgtgtgatcagcttgccttatatctcgtctgaagaagcgcata 1445
 QY 394 tcgacatactgacgagcttcgacatcgacacgacgacgacgacgacgacgacgacgacgac 453
 DB 1446 ttgatatactatgtatccttcctatcgtatcgtatcgtatcgtatcgtatcgtatcgtatc 1505
 QY 454 ctaagacaaatcattatgtatccttcgagacacgacgacgacgacgacgacgacgacgacgac 513
 DB 1506 agaaacacaaatcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1565
 QY 514 agctgcagttcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 573
 DB 1566 aggcacacgtacgcgcggtgtcgttcaagaattcgtcgaattcgtcgaattcgtcgaattcgtc 1625
 QY 574 acgagctcaccagcgcccggtgtgatagccggtgtgaagaggtgtgcaaacctgctcacc 633

DB 1626 accgtgtcacccggtcaggtatcgtcacaaggggttgtgaagagcgtcacaggaaccacg 1685
 QY 634 aggaaccacgggggttgcgtatcgtcagacgtcctctctcctcctcctcctcctcctcctcctc 693
 DB 1686 atcagccaaagagggcgtgtgtatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1745
 QY 694 gagacatacccgcggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 753
 DB 1746 ggaacatacccgagagagacggtgtgagatgtgcacaaacgtacataagacgttctgtatgtgat 1805
 QY 754 tcatcgcagcagctcagacatgt 813
 DB 1806 tcatcgcagcagacagacatgt 1865
 QY 814 cgggggtgtgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 873
 DB 1866 caggtcgtggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1925
 QY 874 aggtcgtcagcagcgtacacgtatcgtatcgtatcgtatcgtatcgtatcgtatcgtatcgtatc 933
 DB 1926 aggtcgtcagcgt 1985
 QY 934 gagaccacaaagtcgaaggtgtcgcgtcaccgcaagcggtgtggaagctta 935
 DB 1986 gagatccaaacagtggaaggtgtgagcgttataagaagaacgagcgtgtggtatgtccta 2037

RESULT 3
 V31505
 ID V31505 standard; DNA: 2330 BP.
 AC V31505;
 XX
 XX 14-AUG-1998 (first entry)
 DT
 DE Base sequence of DNA fragment containing yeast URA3 gene.
 XX
 XX Candida utilis; yeast vector; promoter; marker gene; GAP gene;
 KW high-efficiency integration; monellin; food; drug; L41; URA3; ss.
 XX
 OS Candida utilis.
 XX
 FH Key Location/Qualifiers
 FT CDS 1259..2062
 FT /*tag= a
 FT /note= "URA3 gene segment"

PN MO9807873-A1.
 XX
 XX 26-FEB-1998.
 PD
 XX
 XX 22-AUG-1997; 97MO-JP02924.
 PF
 XX
 XX 23-AUG-1996; 96JP-0241062.
 PR
 XX
 XX (KIRI) KIRIN BEER KK.
 PA Kondo K, Miura Y;
 PI
 XX
 XX WPI; 1998-169177/15.
 DR P-PSDB; W57489.
 DR
 XX
 XX Yeast vector for multi-copying on to chromosomes of yeast such as
 PT Candida utilis - contains a shortened promoter sequence linked to a
 PT marker gene for high-efficiency integration
 XX
 PS Claim 28; Fig 5; 107pp; Japanese.
 XX
 CC This represents the base sequence of DNA fragment containing yeast URA3
 CC gene. This can be used in the construction of a yeast vector for
 CC multi-copying on to chromosomes of yeast such as Candida utilis. The
 CC vector contains a DNA homologous with a chromosomal gene of the yeast

CC (preferably ribosomal DNA (rDNA) such as the URA3, LAI, PGK, GAP or PMA
 CC gene), a marker gene to be used in transformant selection (such as a
 CC drug resistance gene, e.g. the cycloneximide resistance gene LAI, the
 CC 6418 resistance gene Th903-APV, or the hygromycin B resistance gene
 CC (from E.coli) HPT), a shortened promoter sequence such as the C.utillis
 CC LAI, phosphoglycerate kinase (PGK), glyceraldehyde-3-phosphate
 CC dihydrogenase (GAP) or plasma membrane proton ATPase (PMA) gene promoter
 CC and a gene of interest from yeast or other origin. The vectors are useful
 CC in the preparation of proteins for food or drug use in high efficiency.
 CC They can be used especially for the production of single-chain monellin,
 CC which is a low-calorie sweetener whose thermostability is greater than
 CC that of the dimeric natural monellin. Multiple copies of the desired gene
 CC are integrated into the yeast chromosome and high expression efficiency
 CC is obtained.

XX Sequence 2330 BP; 669 A; 498 C; 613 G; 550 T; 0 other;

Query Match 26.8%; Score 370.4; DB 19; Length 2330;
 Best Local Similarity 67.5%; Pred. No. 1,9e-96;
 Matches 531; Conservative 0; Mismatches 251; Indels 0; Gaps 0;

QY 214 caacgaatctctacgagaagaaggccaacatctgcagctgctagaagcttc 273
 DB 1266 cagctatctgacacagagagagcagcagcccttcgacactgtgaagcttct 1325
 QY 274 tggcatctgacagagaagaacacatctctgctcctctatgtgaggaagctta 333
 DB 1326 tctcgctatctgagatcagaagaagacacctgtgtccagtgctgagcttcacag 1385
 QY 334 gaagatctctgagctagacagacagcttgagacacattgctcgcaggaacatct 393
 DB 1386 aggaagcttgcagacagctgtgataagctgtgctctatctcgtctgttgaagaagcata 1445
 QY 394 tgcacatactgacagactctgcacatcgaacagacagcttcagcagcagcttgcg 453
 DB 1446 tctgatactgatactctctcctcctcctgagcttgcctcactgttgcgcttcaa 1505
 QY 454 ctaagacacactcagatctctcgaagacagcaagcttgcctgaacttgcagacagctta 513
 DB 1506 agaaagacacactctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1565
 QY 514 agctgcagtaactctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 573
 DB 1566 aggcacagctacgcgcgtgtgtcgtcctcctcctcctcctcctcctcctcctcctc 1625
 QY 574 acggcgtcacgcgcgcgtgtgtatagccggtcgaagagctcgcgaactgtgcctcac 633
 DB 1626 acggtgtcacgcgcgtgtgtatagccggtcgaagagctcgcgaactgtgcctcac 1685
 QY 634 aggaacccaggggtgtgtcgtatgctgtcgaagagctcctcctcctcctcctcctcctc 693
 DB 1686 atgagccaaagaggtgtgtatgctgtcgaagagctcctcctcctcctcctcctcctc 1745
 QY 694 gaacacataccgcgcgcgtgtgtatgaaatgacgaagctgtgtaacatctgtaagct 753
 DB 1746 ggcacataccgcgcgcgtgtgtatgaaatgacgaagctgtgtaacatctgtaagct 1805
 QY 754 tcatcgacgcgcgtgtgtatgaaatgacgaagctgtgtaacatctgtaagct 813
 DB 1806 tcatcgacgcgcgtgtgtatgaaatgacgaagctgtgtaacatctgtaagct 1865
 QY 814 cgggggtgtgtcgtcgcgcgcgtgtgtatgaaatgacgaagctgtgtaacatctgtaag 873
 DB 1866 caggcgtgtgtcgtcgcgcgcgtgtgtatgaaatgacgaagctgtgtaacatctgtaag 1925
 QY 874 aggtctcgcgcgcgtgtgtatgaaatgacgaagctgtgtaacatctgtaagct 933
 DB 1926 aggtctcgcgcgcgtgtgtatgaaatgacgaagctgtgtaacatctgtaagct 985
 QY 934 gaagaccacaggtgtgtatgaaatgacgaagctgtgtaacatctgtaagct 1045
 DB 1986 gaagaccacaggtgtgtatgaaatgacgaagctgtgtaacatctgtaagct 2037

RESULT 4
 Q66476
 ID 066476 standard; DNA; 1469 BP.
 XX
 AC 066476;
 XX
 DT 17-FEB-1995 (first entry)
 XX
 DE K. marxianus URA3 gene.
 XX
 KW Polymerase chain reaction; primer; amplify; PCR; extension assay;
 KW functionality; inulinase; promoter; regulatory region; promoter;
 KW upstream activating sequence; upstream repressing sequence; UAS;
 KW terminator; secretory signal; secretion; yeast; transformation;
 KW protein; metabolite; ss.
 XX
 OS Kluyveromyces marxianus.
 XX
 FH Key Location/Qualifiers
 FT CDS 546..1349
 FT /*tag= a

W09413821-A.
 XX
 PD 23-JUN-1994.
 XX
 PF 09-DEC-1993; 93WO-EP03547.
 XX
 PR 11-DEC-1992; 92EP-0203932.
 PR 29-JAN-1993; 93EP-0200240.
 XX
 PA (UNIL) QUEST INT BV.
 XX
 PI Chapman JW, Musters W, Rouwenhorst RJ, Toschka HX, Verbakel JMA;
 XX
 DR WPI: 1994-217894/26.
 DR P-PSDB: R56259.
 XX
 PT Yeast regulatory sequence from the inulinase gene - for protein,
 PT RNA or metabolite prodn.
 XX
 PS Example 13; Fig 12; 68pp; English.

XX This sequence represents URA3 gene from K. marxianus. This gene was
 CC used in an example in the construction of the uracil requiring K.
 CC anus mutants, KMS3. This gene is involved in the regulation of the
 CC biosynthesis of pyrimidine. Yeast cells which have an active URA3
 CC gene are unable to grow in medium containing 5-fluoro-orotic acid
 CC (5-FOA). KMS3 is a non-reverting K. marxianus leu2 strain which is
 CC suitable for overexpression of homologous or heterologous proteins in
 CC the food industry.
 XX
 SQ Sequence 1469 BP; 443 A; 258 C; 356 G; 412 T; 0 other;

Query Match 26.8%; Score 369.6; DB 15; Length 1469;
 Best Local Similarity 66.1%; Pred. No. 2.6e-96;
 Matches 534; Conservative 0; Mismatches 274; Indels 0; Gaps 0;

QY 194 accatgcatgacacatgtcaacgaatcttaacgagaaggccaagcacacatctc 253
 DB 530 aactaatcaagaagagatgtcgtactaagatctactcgaagaagcagctcctcatgaag 589
 QY 254 gccagttgctagaagactctcgtcatgtgacagagaagaacaacatctcgtgcttc 313
 DB 590 tccagttgctcgaactcttaactgtatgagaagaagaagtcacaaactatgtgttc 649
 QY 314 ccttgatgttgagcgcgtcgtgaagctctcgtgagctgacagacagcttggaacagc 373
 DB 650 tcttgatgttgtaaaacagcagagctgttaagaatgtcgtgagcttcttggtccatata 709

PT veterinary medicine, for treatment of existing infections and for
 PT prevention of these in immune depressed subjects

XX
 PS Claim 34: Page 53-58; 76pp; German.

CC This sequence is plasmid PK9/10 which is used in a method for the
 CC identification of antimicrobial agents using as a target a nucleic acid
 CC which controls an essential protein of *Saccharomyces cerevisiae* or
 CC from other species of *Mycetes*. Such agents are potentially useful
 CC clinically, in human or veterinary medicine, for treating existing
 CC infections and for preventing them in immune-depressed subjects (those
 CC with human immune deficiency virus infection or diabetes), also as
 CC fungicides and preservatives for foods and body care products. The agents
 CC are used to identify equivalent genes in other fungi, specifically
 CC *Candida albicans* or *Aspergillus fumigatus*, and equivalent human, animal
 CC and plant genes, and also for identification of antimicrobial agents.

XX
 SQ Sequence 4102 BP; 1111 A; 924 C; 1020 G; 1047 T; 0 other;

Query Match 25.8% Score 356; DB 19; Length 4102;
 Best Local Similarity 65.5%; Pred. No. 3.4e-92;

Matches 521; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

OY 209 catgtcaacgaattctacgcagaagggcacaagcacaacattcgccagttgtagaag 268
 DB 935 catgtcgaagctacataaagaacgctgtctactactctcctagctcgtgtgcgca 994
 OY 269 gctctcgcattgacgagagaagaaacaaatctcgtcgtcctcccttgatgtcgag 328
 DB 995 gctatttaatacagcagaagaaacaaactgtgtgtctcatgtgattcgtctac 1054
 OY 329 gtctagaagctctctgtgagctagcagacacgcctggacgcacattctgtcgtaaag 388
 DB 1055 caccagaagattactcgtgagcttagtgaaagcattagtcaccaaatcttgtaactaaac 1114
 OY 389 acatgtcgcacatacgcaggaacttcgacatcgagacagcagtcagccgtcgcagagct 448
 DB 1115 acatgtcgcacatacgcaggaacttcgacatcgagacagcagtcagccgtcgaagcatt 1174
 OY 449 tgcgcgtacgaacacactatcgtactcgtcgaagcagcagtcgctgactgtgaaacac 508
 DB 1175 atcgcgcgaagttcaatttttctcctcgcgaagcagaacaaattcgtcgaactgtgaaac 1234
 OY 509 ggttaagctcgaatctcctcgcgcgtgtacacgtacgcggagtggtggagatataccaa 568
 DB 1235 agtcaaatgcgtactcctcgcgcgtgtacacgtacgcggagtggtggagatataccaa 1294
 OY 569 tgcacagcgcgtcaccgc 628
 DB 1295 tgcacagcgcgtcaccgc 1354
 OY 629 ctcaacagaacacagggggtgtgtcgtatgtcgcagagctcctcctcgcgcgcgcgcgc 688
 DB 1355 aacaaagaacacagggggtgtgtcgtatgtcgcagagctcctcctcgcgcgcgcgcgc 1414
 OY 689 ggcgcggaactatacgc 748
 DB 1415 tacttgagaataatacgaaggttactgtgacattcgcgaagcgcgaacaaagtttgttat 1474
 OY 749 cgggttctacgc 808
 DB 1475 cgggttctacgc 1534
 OY 809 gaccgc 868
 DB 1535 gacacccgc 1594
 OY 869 ggtatgaggtcgc 928
 DB 1595 ggtatgaggtcgc 1654
 OY 929 gggagaagacacccaaagtcgaggggtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 988

DB 1655 gggagaagacacccaaagtcgaggggtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1714
 OY 989 ggc 1004
 DB 1715 gagaagatcgcgccag 1730

RESULT 9

ID T04575 standard; DNA; 4164 BP.
 XX
 XX T04575;
 AC
 XX 30-JAN-1996 (first entry)
 DT
 XX
 DE Plasmid PAT-1 sequence.

KW PAT-1; plasmid; artificial transposon; DNA sequencing; mapping;
 KW genome; ds; cyclic.

OS Chimeric *Saccharomyces cerevisiae*;
 OS Chimeric *Escherichia coli*;
 OS Chimeric synthetic.

PN WO9523875-A1.

PD 08-SEP-1995.

PE 02-MAR-1995; 95MO-US02520.

PR 02-MAR-1994; 94US-0204675.

PA (UYJO) UNIV JOHNS HOPKINS.

PI Boeke JD, Bratlerman LT, Devine SE;

XX WPI; 1995-320587/41.

PT In vitro transposition of artificial transposon(s) - useful e.g. for
 PT sequencing entire genome(s) of particular organisms

PS Disclosure; Page 44-46; 83pp; English.

XX PAT-1 (PSD544) includes: the yeast URA3 gene; a bacterial origin of
 CC replication and multicloning site; a selectable marker gene; 2
 CC blunt-ended transposon termini (substrates for yeast retrotransposon
 CC Ty1 integrase) which flank a restriction site useful for insertion of
 CC a second selectable marker gene; and a second restriction site that
 CC flanks the transposon termini. Digestion of PAT-1 at the second
 CC restriction site liberates an artificial transposon useful for DNA
 CC sequencing and genome mapping.

SQ Sequence 4164 BP; 1185 A; 886 C; 988 G; 1105 T; 0 other;

Query Match 25.8% Score 356; DB 16; Length 4164;
 Best Local Similarity 65.5%; Pred. No. 3.5e-92;

Matches 521; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

OY 209 catgtcaacgaattctacgcagaagggcacaagcacaacattcgccagttgtagaag 268
 DB 416 catgtcgaagctacataaagaacgctgtctactactcctagctcgtgtgcgca 475
 OY 269 gctctcgcattgacgagagaagaaacaaatctcgtcgtcctcccttgatgtcgagac 328
 DB 476 gctcatataatcagcagaagaaacaaactgtgtgtctcatgtgattcgtctacac 535
 OY 329 gtctagaagctctcgtgagcttagcagacacgcctgggagccgcacattgtctgtagaag 388
 DB 536 caccagaagattactcgtgaggttactgtgaaagcattagtcaccaaatcttgtaactaaac 595
 OY 389 acatgtcgcacatacgcaggaacttcgacatcgagacagcagtcacagcgcgtcgcagagct 448


```

Dh      896  tacggagaatactaaaggtactgtttgacattcgaaagacgacaagaattttgtat 955
Oy      749  cgggttcacgcgcagcgtgcacatggtggtggtggtgcgcagcggcttgaactgcacatc 808
Db      956  cggcttatctgcacaagagacacatggtgtggaagagatgaaggttaagatctgtgtatc 1015
Oy      809  gaccgcgggggtttgacctgcagacccaagaagagacgacctgtggccagcagatcacgcacgt 868
Db      1016  gacacccgggtgtggtgtttaga tgcacaaggagagacgcatalgtgtcaacagatataaccgt 1075
Oy      869  ggaatgaggtctgcacgcgcgtacccagatgtgacatcatgtgtgcagaagggctcttgacaa 928
Db      1076  ggaatgattgtgtctcttaagagatctgacatataatcatgtgtggaagagactatttgcaaa 1135
Oy      929  gggaaagagaccacaaggtcgaagggtgcccgcgtacgcgaagcgcgtgtggagaagcttact 988
Db      1136  gggaaagagatcgtcaagtagaggtgtaacgttcaagaagaacagcgtcgggaagacatattc 1195
Oy      989  gcgcgcgtatgggcgag 1004
Db      1196  gagaagatcgcggccag 1211

RESULT  11
v69740
ID      V69740 standard; DNA: 4164 BP.
XX
AC      V69740;
XX
DT      04-FEB-1999 (first entry)
XX
XX      Nucleotide sequence of artificial transposon AT-1 (PAT-1) plasmid.
DE
XX
KW      Transposon; artificial; marker gene; U3; yeast; retrotransposon;
KM      Tyl; DNA mapping; DNA sequencing; primer site; PAT-1; ds.
XX
OS      Synthetic.
OS      Saccharomyces cerevisiae.
XX
PN      US5843772-A.
XX
PD      01-DEC-1998.
XX
XX      10-JUN-1996; 96US-0660754.
PF
XX
PR      02-MAR-1995; 95US-0397679.
PR      02-MAR-1994; 94US-0204675.
PR      10-JUN-1996; 96US-0660754.
PA
XX      (UVCJO ) UNIV JOHNS HOPKINS.
XX
XX      Boeke JD, Bralerman LT, Devine SE;
XX      WPI: 1999-044590/04.
XX
XX      Artificial transposons with terminal Tyl U3 sequences - useful for
PT      DNA sequencing
XX
XX      Examples; Fig 8A-C; 37pp; English.
XX
XX      This represents the nucleotide sequence of an artificial transposon AT-1
CC      plasmid (PAT-1). The invention provides methods for creating artificial
CC      transposons and inserting these transposons into DNA targets in vitro.
CC      The artificial transposon consists of a linear blunt-ended DNA molecule
CC      comprising a marker gene flanked by terminal sequences, each comprising
CC      at least 4 bp of a U3 sequence of yeast retrotransposon Tyl1. The DNA
CC      molecule used for generating such artificial transposons comprise an
CC      origin of replication; a first selectable marker gene; two blunt-ended
CC      transposon terminal, each of at least 4 bp, that are substrates for yeast
CC      retrotransposon Tyl1 integrase and flank a first restriction site useful
CC      for insertion of a second selectable marker gene to form an artificial
CC      transposon, and a second restriction site flanking the two transposon

```

CC terminal where digestion with the second restriction enzyme liberates
CC blunt-ended fragment having the transposon terminal at either end, the
CC fragment being an artificial transposon. The method is useful for DNA
CC mapping and sequencing of e.g., both functional and non-functional
CC sequences, primer sites and restriction sites.

SQ Sequence 4164 BP; 1185 A; 886 C; 988 G; 1105 T; 0 other;

Sequence 4164 BP; 1185 A; 886 C; 988 G; 1105 T; 0 other;

Query Match	25.8%;	Score 356;	DB 20;	Length 4164;
Best Local Similarity	65.5%;	Pred. No. 3.5e-92;		
Matches 521; Conservative	0;	Mismatches 275;	Indels 0;	Gaps 0

QY	209	catgtcaacggaattcttaacgcgaagaagggcccaagcgacacaatttcgcagatttgcctagaaga	268
Db	416	catgtcgaagaagctacacataataagaagctgtctgtactcaactccagtacctgtctgtcgcga	475
QY	269	gtctctgtgcatttgaatgcgagagaagaacaaacctctctgtgcgtctcccttgatgtgtgcgac	328
Db	476	gctatttaattatctatgcacgaanaagacaaacaaactgtgtgtctatctatgtgtatctgtac	535
QY	329	gtctagaagaagctcttgtagcttaacgacacacacgctcggaagccgcagctatttctgtcgtagaag	388
Db	536	caccaaggaattactgtgagttaagtctgaagcataggctcccaaatgtttactctaaaac	595
QY	389	acatgtcgacaacttgcagcgaacttcgacacatcgagacagcaagctcaagccgtctgcagcagct	448
Db	596	acatgtgatatactctgtgcatttctccatgtgagcgacagcttaagccgctaaagcat	655
QY	449	tgcggtctgaagacacaacttcatlgatcttctcgagagccgcgaagtgctgcgtcttgacatttgcgaac	508
Db	656	atccgcacgaatcacatcttctactctcgagagccagaaaaattctgtgcacattgtgtataac	715
QY	509	ggttaagctgtgcagtaacctctccgcgctgtcatccgtatctcgagagctgtgtgcggaattatccaa	568
Db	716	agtcacaattgtgaatcaccctctgcggtgtgtatcacagataagcagaatgtgtgcagaacttcgaa	775
QY	569	tgacacagcgtctcaacgcgcgcgcgtgtgtatagccggtctgaagagcgtctgcgaactgtgc	628
Db	776	tgacacagctgtgtgtgtgcgcacagctaatgtgttaagcgttctgaagcagcgtcgacagaagaagt	835
QY	629	ctcacaggaacacccagggggtgtctctgaatgcctgcgcgaagcctctctctccaggtctcttctgtgc	688
Db	836	aacaaaggaacacttagaagcctcttgaatgttaagacagaattgtcatgtcaaaaggtccctctc	895
QY	689	gctgcgagagactataccgcgcggtgcgtctgtgaatactgcgaaagctgcgaaagactgtgtgat	748
Db	896	tactgtgagaataatactcaaggttaactgtgtgcattctgtgaagagcgcgaacaaagtattgttat	955
QY	749	cgggttctacgcgcagcgtgtgacatgtgtgtgtgtgcgtgtgcagacggtcttgatctgtccatcat	808
Db	956	cggctcttctctcaaaagaaacatgtgtgtgaagaaagctgaaggtctacagattgtgttatat	1015
QY	809	gaaccccggtgtgtgtgacctctgtaacgaacaaaggaagacgcgcgtgcgcagcagcttaccggaagct	868
Db	1016	gacaccccggtgtgtgtgttcttaagatgtcaagggagagacgtgaggttccaaacagtataagaacct	1075
QY	869	ggaatgaagctgcgtcgcgcgcgtacacgaatgtgatacatgtctgtgcagagaggtctctcttgacaa	928
Db	1076	ggatgtatgtgtgtctctcaacagatcttgacattatattgtctggaagaggaactattgtgcaaa	1135
QY	929	ggggaagagaccccaaggtctgaaggtgtcccgctctacccgcaagagcccggttgtgaagcttactt	988
Db	1136	gggaagaggaatctctaagctaaaggttgaaacgcttaacagaagaagcagcgtctggaagacatactt	1195
QY	989	gcgcgcgtatgtgcgag 1004	
Db	1196	gagaagaatgcgcgcaag 1211	

RESULT	12
X60602	
ID	X60602 standard; DNA; 4292 BP

CC PAT-2 (PSD545) includes: the yeast URA3 gene; a bacterial origin of
 CC replication and multicloning site; a selectable marker (DHFR) gene; 2
 CC blunt-ended transposon termini (substrates for yeast retrotransposon
 CC Tyl integrase) which flank a restriction site useful for insertion of
 CC a second selectable marker gene; and an XmnI site that flanks the
 CC transposon termini. Digestion of PAT-2 at the XmnI site liberates
 CC a primer island artificial transposon (PATr, see T04570) useful for
 CC DNA sequencing and genome mapping.

XX Sequence 4933 BP: 1431 A; 1031 C; 1157 G; 1314 T; 0 other:

Query Match 25.8%; Score 356; DB 16; Length 4933;
 Best Local Similarity 65.5%; Pred. No. 3.7e-92;
 Matches 521; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

209 catgtcaacgaatacttcacgagaagggcacaacattccagctgtctagaaa 268
 416 catgtcgaagaagctataaagaaagctgtctactcactcctgtctgtccaa 475
 269 gcttcgacatgacgacgagagaagaacacatctcgtcctccctgtatgtcgag 328
 476 gctatttaatacatcagcagaagaaacaaactgtgtctcattcattgtctgtac 535
 329 gctcgaagaagctctcgaagctagcagacacgcttgagacgacatttctcgtcgaagac 388
 536 caccagaagaattactgaggttaggtgaagcattagctcccaaatctgttactcaaaac 595
 389 aatctgacatactgacgagacttcgacatcgaagacagcagcctgtcagagct 448
 596 aatctgacatactgacgagacttcgacatcgaagacagcagcctgtcagagct 448
 449 tgcgctcgaagacacacttcagatctcgaagacgagcttcgacatctgtcgaacac 508
 656 atccgccaagatacaatttctactcttcgaagacagaaattctgtcactgtgtaaac 715
 509 ggttaagctcgaactactcctccggtcggtgtaaccgatacgcgagtgagataacaa 568
 716 agtcaaatctgacgactctcgtggtgtaacagatagcagaatctgtcagacatctaa 775
 569 tgcagaagcgtctcaccggtcccggtgtgataagcggcgtgaagaagctcgaactggc 628
 776 tgcacacggtctgtgtggtccaggtatctgtcagcggttgcgaagcagcgcaagaagat 835
 629 ctcaacgaagaccaggggtgtgtcgtgactgtcgaagcctcctcgaaggtccttggc 688
 836 aacaagaagaacctgaagagccttctgtatgttagcagaattgtcagcaagggtccctatc 895
 689 ggcggggaacatactccggtcggtcgttgaatgtgcaagctggaagcagacttggat 748
 896 tactggaataatactcaagggtactgtgtacatctgcgaagacgacaaagatctgtat 955
 749 cgggtctcagcagcgttgaacatgagtggtcggtcgaagcgtcttgaagcgtccatc 808
 956 cggctctatctgtcacaagaagacatgggtgtgagaagatgaaagtttaagatgtgtat 1015
 809 gaacccgggtgtgtgtcgtcgaagacaaagagcagcgttggcgaagcagatccgacgt 868
 1016 gaacccgggtgtgtgtcgtcgaagacaaagagcagcgttggcgaagcagatccgacgt 868
 869 ggaagagctcagcagcgttgaacatgagtggtgtgtgtgtgtgtgtgtgtgtgtgtgt 928
 1076 ggt 1135
 929 gggagaagaccaccaaggtcgaaggtgcgcgttaccgaagcgggttggagagctactt 988
 1136 gggagaaggtcgtcgtcgaaggtcgaaggttaccgaagcgggttggagagctactt 1195
 989 ggcgcgt 1004
 1196 gagaagatcgcgcgaag 1211

RESULT 14

V22272

ID V22272 standard; DNA; 4933 BP.

XX V22272;

AC V22272;

XX 07-JUL-1998 (first entry)

XX PAT-2 (PSD545).

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

XX Target DNA sequence determination; PAT-2; PSD545; circular;

QY 269 gcttcgtgacttgatgacgagagaagaaccatctcgtgctcccttgatgtcgagac 328
 DB 476 gctatttaatatatgacgaagaagaacaacttgctgctcattggtatcgtaac 535
 QY 329 gctcagaagaactctcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 388
 DB 536 caccagaagaaatctcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 555
 QY 389 acatgtcgaacatactcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 448
 DB 596 acatgtgatactcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 655
 QY 449 tggggtcgaacacacactcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 508
 DB 656 atccgcaagtaacatacttctcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 715
 QY 509 ggttaagctcgaactcctcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 568
 DB 716 agtcaaatctcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 775
 QY 569 tgcacacggtcgtcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 628
 DB 776 tgcacacggtcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 835
 QY 629 ctcaacaggaacccaggggtgtgctgagctagacagacacgctggagaccgacatctgtctgtagaac 688
 DB 836 aacaaaggaacccaggggtgtgctgagctagacagacacgctggagaccgacatctgtctgtagaac 895
 QY 689 ggcgcgaggaacatactcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 748
 DB 896 tactgaggaatactcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 955
 QY 749 cgggttcacatcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 808
 DB 956 cgggttcacatcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 1015
 QY 809 gaccgcgaggtcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 868
 DB 1016 gacacccggtggtgtgagctagacagacacgctggagaccgacatctgtctgtagaac 1075
 QY 869 ggaatgagctcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 928
 DB 1076 ggaatgagctcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 1135
 QY 929 gggaaaggaacccaggggtgtgctgagctagacagacacgctggagaccgacatctgtctgtagaac 988
 DB 1136 gggaaaggaacccaggggtgtgctgagctagacagacacgctggagaccgacatctgtctgtagaac 1195
 QY 989 ggcgcgtagtgagag 1004
 DB 1196 gagaagatcgagcag 1211
 RESULT 15
 V69741
 ID V69741 standard; DNA: 4933 BP.
 AC V69741;
 DT 04-FEB-1999 (first entry)
 DE Nucleotide sequence of artificial transposon AT-2 (PAT-2) plasmid.
 XX Transposon; artificial; marker gene; U3; yeast; retrotransposon;
 KW Ty1; DNA mapping; DNA sequencing; primer site; PAT-2; ds.
 XX
 OS Synthetic.
 OS Saccharomyces cerevisiae.
 XX
 PN US5843772-A.
 XX
 PD 01-DEC-1998.
 XX

PF 10-JUN-1996; 96US-0660754.
 XX 02-MAR-1995; 95US-0397679.
 PR 02-MAR-1994; 94US-0204675.
 PR 10-JUN-1996; 96US-0660754.
 XX
 FA (U3) UNIV JOHNS HOPKINS.
 XX
 PI Boeke JD, Bralerman LT, Devine SE;
 DR WPI: 1999-044590/04.
 PT Artificial transposons with terminal Ty1 U3 sequences - useful for
 PS DNA sequencing
 PS Examples; Fig 9A-C; 37pp; English.
 CC This represents the nucleotide sequence of an artificial transposon AT-2
 CC plasmid (PAT-2). The invention provides methods for creating artificial
 CC transposons and inserting these transposons into DNA targets in vitro.
 CC The artificial transposon consists of a linear blunt-ended DNA molecule
 CC comprising a marker gene flanked by terminal sequences, each comprising
 CC at least 4 bp of a U3 sequence of yeast retrotransposon Ty1. The DNA
 CC molecule used for generating such artificial transposons comprise an
 CC origin of replication; a first selectable marker gene; two blunt-ended
 CC transposon terminal, each of at least 4 bp, that are substrates for yeast
 CC retrotransposon Ty1 integrase and flank a first restriction site useful
 CC for insertion of a second selectable marker gene to form an artificial
 CC transposon, and a second restriction site flanking the two transposon
 CC terminal, where digestion with the second restriction enzyme liberates a
 CC blunt-ended fragment having the transposon terminal at either end, the
 CC fragment being an artificial transposon. The method is useful for DNA
 CC mapping and sequencing of e.g. both functional and non-functional
 CC sequences, primer sites and restriction sites.
 SO Sequence 4933 BP; 1431 A; 1031 C; 1157 G; 1314 T; 0 other;
 Query Match 25.8%; Score 356; DB 20; Length 4933;
 Best Local Similarity 65.5%; Pred. No. 3.76-92;
 Matches 521; Conservative 0; Mismatches 275; Indels 0; Gaps 0;
 QY 209 catgtcgaacgaatcttcaacgagaaggaaggaacacacatccgacgttctagaaa 268
 DB 416 catgtcgaacgaatcttcaacgagaaggaaggaacacacatccgacgttctagaaa 475
 QY 269 gcttcgtgacttgatgacgagagaagaaccatctcgtgctcccttgatgtcgagac 328
 DB 476 gctatttaatatatgacgaagaagaacaacttgctgctcattggtatcgtaac 535
 QY 329 gctcagaagaactctcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 388
 DB 536 caccagaagaaatctcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 555
 QY 389 acatgtcgaacatactcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 448
 DB 596 acatgtgatactcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 655
 QY 449 tggggtcgaacacacactcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 508
 DB 656 atccgcaagtaacatacttctcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 715
 QY 509 ggttaagctcgaactcctcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 568
 DB 716 agtcaaatctcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 775
 QY 569 tgcacacggtcgtcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 628
 DB 776 tgcacacggtcgtcgtgagctagacagacacgctggagaccgacatctgtctgtagaac 835
 QY 629 ctcaacaggaacccaggggtgtgctgagctagacagacacgctggagaccgacatctgtctgtagaac 688
 DB 836 aacaaaggaacccaggggtgtgctgagctagacagacacgctggagaccgacatctgtctgtagaac 895

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2001, 21:40:18 ; Search time 81.04 Seconds

(Without alignments)
2744.339 Million cell updates/sec

Title: US-09-582-779A-1

Perfect score: 1380

Sequence: 1 cccgagcaccatcattggaag.....aatatgatacgcttgaaattc 1380

Scoring table: IDENTITY_NUC

Searched: 280836 segs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgnt2_6/prodata/1/lna/5A_COMB.seq:*

2: /cgnt2_6/prodata/1/lna/5B_COMB.seq:*

3: /cgnt2_6/prodata/1/lna/6_COMB.seq:*

4: /cgnt2_6/prodata/1/lna/PCTUS_COMB.seq:*

5: /cgnt2_6/prodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	370.4	26.8	2330	2	US-08-557-128-3
2	357	25.9	5181	3	US-08-801-344-5
3	356	25.8	4164	1	US-08-204-675-1
4	356	25.8	4164	2	US-08-660-754-1
5	356	25.8	4164	2	US-08-796-364-1
6	356	25.8	4164	4	PCT-US95-02520-1
7	356	25.8	4933	1	US-08-204-675-2
8	356	25.8	4933	2	US-08-660-754-2
9	356	25.8	4933	2	US-08-796-364-2
10	356	25.8	4933	4	PCT-US95-02520-2
11	356	25.8	7102	3	US-09-138-024-20
12	356	25.8	7333	3	US-09-138-024-21
13	356	25.8	7633	3	US-09-028-851-1
14	356	25.8	7633	3	US-08-815-520-1
15	352.8	25.6	1115	1	US-08-507-455-5
16	349	25.3	1220	5	5204252-1
17	336.4	24.4	906	5	5204252-3
18	289.4	21.0	2688	1	US-08-088-633-3
19	289.4	21.0	2688	1	US-08-245-756-3
20	289.4	21.0	2688	1	US-08-441-750-3
21	289.4	21.0	2688	4	US-08-441-751-3
22	289.4	21.0	2688	4	PCT-US92-02521-3
23	115.2	8.3	8533	1	US-07-845-181-6
24	115.2	8.3	8533	1	US-07-845-989-6
25	39.4	2.9	7218	1	US-08-232-463-14
26	38.4	2.8	1931	2	US-09-130-114-2
27	38.2	2.8	2779	3	US-08-482-677-5
28	38.2	2.8	2783	1	US-08-152-019A-41

C 29	36.6	2.7	43280	2	US-08-804-227C-1	Sequence 1, Appl1
C 30	36	2.6	24417	2	US-08-846-762-1	Sequence 1, Appl1
C 31	35.8	2.6	20235	3	US-07-642-734C-3	Sequence 3, Appl1
C 32	35.8	2.6	20235	3	US-08-439-009A-3	Sequence 3, Appl1
C 33	35.4	2.6	1182	1	US-08-241-943-23	Sequence 23, Appl1
C 34	35.4	2.6	1431	1	US-08-254-357-1	Sequence 1, Appl1
C 35	35.4	2.6	2327	5	5512669-5	Patent No. 5512669
C 36	35.4	2.6	2327	5	5512669-5	Patent No. 5512669
C 37	35.4	2.6	4983	1	US-08-472-358-1	Sequence 1, Appl1
C 38	35.4	2.6	4983	4	PCT-US92-0576A-1	Sequence 1, Appl1
C 39	35.4	2.6	4984	1	US-08-687-806-1	Sequence 1, Appl1
C 40	35.2	2.6	3061	1	US-08-700-576-1	Sequence 1, Appl1
C 41	34.4	2.5	28958	1	US-08-258-261B-6	Sequence 6, Appl1
C 42	34.4	2.5	28958	1	US-08-456-837-6	Sequence 6, Appl1
C 43	34.4	2.5	28958	1	US-08-457-342-6	Sequence 6, Appl1
C 44	34.4	2.5	28958	1	US-08-457-646A-6	Sequence 6, Appl1
C 45	34.4	2.5	28958	1	US-08-458-076A-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-08-557-128-3
Sequence 3, Application US/08557128
Patent No. 5849524
GENERAL INFORMATION:
APPLICANT: KONDO, Keiji
APPLICANT: KAWARA, Susumu
TITLE OF INVENTION: TRANSFORMATION SYSTEMS FOR THE YEAST
TITLE OF INVENTION: CANDIDA UTILIS AND THE EXPRESSION OF HETEROLOGOUS GENES
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557,128
FILING DATE: 25-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP95/01005
FILING DATE: 25-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-129287
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-285823
FILING DATE: 26-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-135015
FILING DATE: 25-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/108
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2330 base pairs

```

: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1259..2059
: FEATURE:
: NAME/KEY: mat.peptide
: LOCATION: 1259..2059
: US-08-557-128-3

```

```

Query Match 26.8%; Score 370.4; DB 2; Length 2330;
Best Local Similarity 67.5%; Pred. No. 3.3e-101;
Matches 521; Conservative 0; Mismatches 251; Indels 0; Gaps 0;

```

```

QY 214 caacgaatcttaccgaagaagcgcaaacgcaacatctccagcttctagaagcttc 273
DB 1266 CCACGTTATCGTACACAGAGACGCGATCGACACCCCTCCGCACTGCTAAGCGTCTGT 1325
QY 274 tggcatttgacgagagaagaaccacatctcgcgttcccttgaatgctggagactta 333
DB 1326 TTTCCCTTATGAGTCCAAAGAGACGACCTGTGTGCCAGTGTGCTGTTGACACAG 1385
QY 334 gaaagctcttgagcttagcagacgagcgtggagccgacatctgtctgctgagacacatg 393
DB 1386 AGGAGTGTCTCAAGCTCGTTGATACGCTTGCTTATATCTGTCTGTGAAGACGCTTA 1445
QY 394 tgcacatactgacgagacttcgacatcgagacgacagctcaagcgttgacagacgttcg 453
DB 1446 TTGATATATCTGATGACTCTCTATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1505
QY 454 ctaagacacatctatgactcttcgagagacgcaagctcgcgagacttggcaacagctta 513
DB 1506 ACAAGCACAAATTTCTCATCTTTGAGGACCGTAAAGTTGCTGATATGCGCAACACGCTCA 1565
QY 514 agctcgagtaacctccctccgctgctacgctatcgagtgagtgagcagatcttaccatgac 573
DB 1566 AGGCACATACCGCCGGTGTGCTTACAGATTGCCCATGGCAGATATACACCAACGCC 1625
QY 574 acgagctacgagcccgctgtgataccgagctgaagaagcttgacgaactgagcttacc 633
DB 1626 ACGGTGTACACGGTGCAGGTATGTCAGAGCGTGAAGAGGCTGCACAGAAACACAG 1685
QY 634 aggaacacgggggtgtgctgagctgagagagctctctctcagagggcctcttgagcg 693
DB 1686 ATGAGCAAGAGAGGCTGTGATGCTTGCGAGCTGAGCTCAAGGGCTCTTGCGCCACG 1745
QY 694 ggaacataccgagcgctgctgaatgagcgaagcttgacgaagaacttgtatcgggt 753
DB 1746 GGCATATATACCGAGGAGACCGTGGAGATTGCCAAACATGATAGAGACTTTGTATTGGAT 1805
QY 754 tcatcgagcagcgatgacatgagtgagcgctgacagagcttgaactggtcatcatgaccc 813
DB 1806 TCATCCACAGAGACGATGCTGCGACAGAGATGGGTTGCTGATCATCATGATGATGATG 1865
QY 814 cgggggttgcttgacgacgaagaagagagcgttgagcagagatgacgacagcggtgag 873
DB 1866 CAGGCTGTGGACTCGAGTAAAGGCGAGCTCCCTGGGCCAACACAGAACTGTGCTGAT 1925
QY 874 aggtcgctgagcagcgatgacgagtgatcatgttgcagagagggctctcttgcagaaggaa 933
DB 1926 AGGTTTCTAGTGTGCTGCTGATCATCATCATCATCATCATCATCATCATCATCATCAT 1985
QY 934 gaaaccccaaggtgcagaggtgcgcgtacacgaagcggttgagagagctta 985
DB 1986 GAGATCCAAAGTGAAGAGTGAAGCGTTATGAGAAAGACAGGCTGGATGCTTA 2037

```

```

RESULT 2
US-08-801-344-5/c
; Sequence 5, Application US/08801344
; Patent No. 6087140

```

```

: GENERAL INFORMATION:
: APPLICANT: Cameron, Douglas C.
: APPLICANT: Shaw, Anita J.
: APPLICANT: Alatas, Nedim E.
: TITLE OF INVENTION: MICROBIAL PRODUCTION OF 1,2-PROPANEDIOL
: TITLE OF INVENTION: FROM SUGAR
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dewitt Ross & Stevens S.C.
: STREET: 8000 Excelsior Drive, Suite 401
: CITY: Madison
: STATE: WI
: COUNTRY: U.S.A.
: ZIP: 53717-1914
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/801,344
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Sara, Charles S.
: REFERENCE/DOCKET NUMBER: 09820.037
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 608-831-2100
: TELEFAX: 608-831-2106
: INFORMATION FOR SEO ID NO: 5
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5181 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: MOLECULE TYPE: circular
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Yeast shuttle vector YEP352
: US-08-801-344-5

```

```

Query Match 25.9%; Score 357; DB 3; Length 5181;
Best Local Similarity 62.7%; Pred. No. 5.5e-97;
Matches 555; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

```

```

QY 120 acatcattctgattgtgtagtgcctccatccacagtaagcattgtataagctgtatcc 179
DB 1257 ATATACCATATGTATGTGTAAGAAACATGAATTGCCCCAGTATCTTAACCAACTGC 1198
QY 180 atagggtctacgacacttagcacttgcacatctcaagaactttagcgagaagggcc 239
DB 1197 ACAGAACAAAACCGAAGACGATTAATCATGTGCAAGCTCATATTAAGAACTGCT 1138
QY 240 aagcagacaatttcgcaagcttgctagaagctctctgcatctgaacgagaagaagcc 299
DB 1137 GCTACTATCTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078
QY 300 aatctcgtcctcccttctgctgagagcgtcagaagcttctgagcttagcagaag 359
DB 1077 AACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1018
QY 360 ctggagcgcacatttctcgtgagaaacacatgctgacacttgcagcagctcgacatc 419
DB 1017 TTAGTCCCAAAATTTTACTTAATAAACACATGTGATATCTTGAATGCTGATGCT 958
QY 420 gagaacgaatcgaagccgctgagcagcgtcgagcgtgacacacacttcaatgacttcgag 479
DB 957 GAGGCAACAGTTAAGCCGCTTAAGGCAATTATCCGCAAGTACATTTTACTCTTCGAA 898
QY 480 gaccgcaagctcgtgacatctgacacacagcttaagctcagtaactcctcggcggtgac 539

```



```

Db 656 ATCCGCAAGTACATTTTCTTCTGACAGACGAAATTTGCTGACATTTGGTAATAC 715
Qy 509 ggttaagctcagtaactcctccgagctgtaacgtaacgagagtgagagatattacaa 568
Db 716 AGTCAAAATGAGTACTGCGGGGTGTATACAGATAGCAGATGGGACAGATATACAA 775
Qy 569 tgcacagcgctcaccgcccgcgtgtaagcggcgttgaaagagcgtcgaatactgac 628
Db 776 TGCACACGGGTGTGCGGCGCCAGGTATGTCGCGTTGAAGCAGCGCGGCAAGAGT 835
Qy 629 cccacagagaccagagggctgctgtaacgctcctcctcctcctcctcctcctcctg 688
Db 836 AACAAAGAACCTAGAGGCTTTTGTATGTTACAGATTTGTCATGCAAGGCTCCCTATC 895
Qy 689 ggcgagagactacacccgagcgctgtaaaatggcgaagctggaagacatctgac 748
Db 896 TACTGAGAAATATACAGGTAAGTGTGACATGCGAAGACGACAAAGATTTGTTAT 955
Qy 749 cgggttcacgcgcagcgctgacatgggtgagcgctgacagagccttgactgctacat 808
Db 956 CGGCTTATTTGTCAAAGAGACATGGGTGGAAGAGATGCAAGTTGCTGATTTAT 1015
Qy 809 gaccccgaggtctgagcctgagacaaagagagcgctgagccagctacacgagct 868
Db 1016 GACACCCGGGTGCGTTTGTATGATGACAGAGGAGACGATTTGGTCAACATATAGACCT 1075
Qy 869 ggaatgagctcgtcagcagctacacgctgtaacatctgttgcaagagggctcttgacaa 928
Db 1076 GGATGATGTGTCTCTACAGATCTGACATTTATTTGTTGGAAGGACATTTGCAAA 1135
Qy 929 gggagagagacccagagctgagaggtgagcgcctacacgagcggcttgagagctact 988
Db 1136 GGGAAAGGATCTTAAGGTAGAGGTGAACGTTACAGAAAAAGCAGGCTGGAGACATTTT 1195
Qy 989 ggcgcgtacgagcgag 1004
Db 1196 GAGAGATGCGGCCAG 1211

```

```

TELEFAX: 202.508.9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4933 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: PAT-2
US-08-204-675-2

Query Match 25.8%; Score 356; DB 1; Length 4933;
Best Local Similarity 65.5%; Pfd. No. 1.e-96;
Matches 521; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

Qy 209 catgtcaagaatcctacgagaagggccaaagcagacacacatctgcctccctgagtgcgac 268
Db 416 CATGTCAAAAGCTCATATAGAAAGCGTGTCTACTCATCTGCTGTCTGCTCCAA 475
Qy 269 gctctgacattgacgtcagcagagaagaaacacatctcctgcctccctgagtgcgac 328
Db 476 GCTTATTTATATCTATGCAACGAAAGCAAAACAACTTGTGCTTCAATGATGTTGTTAC 535
Qy 329 gctcagaagctcttgagctgagcagacagcgctgagacgacacattgtctgtaagac 388
Db 536 CACCAAGCAATTTACTGAGGATTTAGTTGAAGCATTTAGTCCCAAAATTTGTTACTAAAC 595
Qy 389 acatgtcagactactgacgactcgcacatcagacagcagcagcagcagcagcagcagc 448
Db 596 ACATGTGATATCTTGAATTTTCCATGAGAGGCAAGTTAAGCGCTAAGAGCATT 655
Qy 449 tgcgctaaagcacacactcactcactcagagagacgagctgcgctgacatctgacacac 508
Db 656 ATCCGCCAAGTACATTTTCTTCTGACAGACGAAAGAAATTTGCTGACATTTGTAATAC 715
Qy 509 ggttaagctcagtaactcctccgagctgtaacgtaacgagagtgagagatattacaa 568
Db 716 AGTCAAAATGAGTACTGCGGGGTGTATACAGATAGCAGATGGGACAGATATACAA 775
Qy 569 tgcacagcgctcaccgcccgcgtgtaagcggcgttgaaagagcgtcgaatactgac 628
Db 776 TGCACACGGGTGTGCGGCGCCAGGTATGTCGCGTTGAAGCAGCGCGGCAAGAGT 835
Qy 629 cccacagagaccagagggctgctgtaacgctcctcctcctcctcctcctcctcctg 688
Db 836 AACAAAGAACCTAGAGGCTTTTGTATGTTACAGATTTGTCATGCAAGGCTCCCTATC 895
Qy 689 ggcgagagactacacccgagcgctgtaaaatggcgaagctggaagacatctgac 748
Db 896 TACTGAGAAATATACAGGTAAGTGTGACATTTATTTGTTGGAAGGACATTTGTTAT 955
Qy 749 cgggttcacgcgcagcgctgacatgggtgagcgctgacagagccttgactgctacat 808
Db 956 CGGCTTATTTGTCAAAGAGACATGGGTGGAAGAGATGCAAGTTGCTGATTTAT 1015
Qy 809 gaccccgaggtctgagcctgagacaaagagagcgctgagccagctacacgagct 868
Db 1016 GACACCCGGGTGCGTTTGTATGATGACAGAGGAGACGATTTGGTCAACATATAGACCT 1075
Qy 869 ggaatgagctcgtcagcagctacacgctgtaacatctgttgcaagagggctcttgacaa 928
Db 1076 GGATGATGTGTCTCTACAGATCTGACATTTATTTGTTGGAAGGACATTTTGCAAA 1135
Qy 929 gggagagagacccagagctgagaggtgagcgcctacacgagcggcttgagagctact 988
Db 1136 GGGAAAGGATCTTAAGGTAGAGGTGAACGTTACAGAAAAAGCAGGCTGGAGACATTTT 1195
Qy 989 ggcgcgtacgagcgag 1004

```



```

Query Match      25.8% ; Score 356 ; DB 3 ; Length 7102;
Best Local Similarity    95.5%; Pred NO. 1,3e-96;
Matches 521; Conservative   0; Mismatches 275; Indels     0; Gaps     0;

QY       209 catgtcaacgaatacttaccgagaagggccaaaggcacacaacttcgcagtcttgtlaganaa 268
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       5814 cagtctcgaagaagcatatcataagaaacgvtgtctactaccttagtccttgctgtcgccaaa 5873
QY       269 gctctcgcattgatgtccaagagaaaaaacatatctctgcgtctccccttgatvgctcgaa 328
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       5874 gcatttaatalcaaagcacagaagaaagcaaaaacttgltgtctctaatltygatvgtttcga 5933
QY       329 gtctaagaagcttlctygaagcttagcaacacacgcgtctyggaacgccscatctgtctcytaa 388
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       5934 caaccaaggatttacitvgagtttagttgaagcatttaggtccccaaatttgyttactttaaaac 5993
QY       389 acatgtcgacatactgacgagcacttcgacalcgagacagatcgaacgcgtcgacayagct 448
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       5994 aaatgtcgatatctctgcagtatcttcccaltgaaagggcacagttaagccgctaaagyacatl 6053
QY       449 tgcggttaagacaacactctatgactcttcgaaggacccgaagctcgcgtcgaattgtgcaacgc 508
          || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       6054 atccgcacaagatacaatttttactctctctcgaagaaacgaanaattctgcgaattgtgcaaac 6113
QY       509 ggtaagctgcagatactctctccgcgctgtaccgylacgcggaagtlyggcggatattaccaaa 568
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       6114 agtcaaatgtcagsgtactctctgtcguggtatagatataagaaatagaaatggcagacattcaga 6173
QY       569 tcacaaagcgctcaacacggcccggtgctgataagccggctcgtaagagggctctcgaaactggc 628
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       6174 tgcaacacggctgctgtggtggcccgagttatgttagcggtcttgaaagcgagcgcgcaagaaga 6233
QY       629 ctcacaggaaacccaaagggggttgctgtatgcctygcagagctctcttccaaaggctctcttggc 688
          || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       6234 aacaaagaacacttagagggcctttigtatgttagcagaattgtcatgcaaaagggtccctcattc 6293
QY       689 gccgagagactataccgcgggcgttcgcttgnaaaltygcgaagcttggaacgaagacttggat 748
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       6294 tacttgaaatatataactlaaagggtacgtctgtgacalctlgcaagagcgacaagaatttgc 6353
QY       749 ccggtctcatcggcagcgctgtaacatgggtgtggcgtgtgcacaacggccttgctgtgtatcat 808
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       6354 cggtcttatctgtcacaagaagacatagggtgtggaaagatgaaagttcagatctgtgtatcat 6413
QY       809 gaaccocgggggtctggcctctgnaagcaaaaagggagacagcgctcttgggccacagataccgacg 868
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       6414 gacaccocggctgtgggtttagatgtacaagggaagacgcatctgggtccaaacagataagaccg 6473
QY       869 gagataagttcgtcaacgcgcgtacccgatcgtgatactatctcttgycacaaggctccttliga 928
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       6474 ggaatgactgtgctcttaccagaagatctgacatatattaactgttgnaaaagagacatttgc 6533
QY       929 gggaagaagaccocaagagctcgaaagtgcccgctcactacgcgaacggcttgaggagcttaact 988
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       6534 gggaaggagatctaaagrtcgagggtggaacggttaacagaaagaacgagcttgggaagcatt 6593
QY       989 ggcgcgctatgggcgag 1004
          || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       6594 gagaagatcgcgccag 6609
          || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT_12
US-09-138-024-21
Sequence 21, Application US/0913802A#
Patent No. 6004779
GENERAL INFORMATION:
APPLICANT: Bradley, John D.
APPLICANT: Thompson, Craig M.
APPLICANT: Moore, Jeffrey B.
APPLICANT: Mobbe, C. Richard
APPLICANT: Healy, Judith M.
APPLICANT: Donnelly, Caroline E.
TITLE OF INVENTION: REGULATED GENE EXPRESSION IN YEAST

```


Db 3468 AGCTAAATTCAGTACCTCTCGGGTGTATACGAAATAGCAGAAATGGGCGACACATTTACGAA 3409.
Qy 569 Tgcacagcgctcaccgagctgtagatagccggcctgaagaagcctcgaacacgac 628
Db 3408 TGCACACGCGTGTGGTGGCCAGGATTTGATGACGCTTTTGAAGCAGCGCGACAGAAAGT 3349
Qy 629 ctcaagaagaccagggggtctgtagatctgtagagctctctcgaagcctcttgac 688
Db 3348 AACAAAGAACCTAGAGCGCTTTTGTATGTTAGCAGAAATTCATGCAAGGCGCTCCATAC 3289
Qy 689 ggcgagagactataccgagcgctgctgtagaagctgtagagcctgaagactctgtag 748
Db 3288 TACTGAGAAATATCTAAGGCTACTGTTGACATTTGCGAAGAGCGCAAAAGATTTGTTAT 3229
Qy 749 cgggttcacgcgacgctgtagatggtggtggtgtagagcgtctgtagagctctgtag 808
Db 3228 CGGCTTTATTTGCTCAAGAGACATGGGTGGAAGAGATGAAAGTTACCATTTGTTAT 3169
Qy 809 gaccccgagggttgccctgtagacagaaagagagcgctgtagcagcagctacgacgct 868
Db 3168 GACACCGGCTGTGGTGTATGATGACAAAGGAGAGCGATTTGGTCAACAGATATGAACCT 3109
Qy 869 ggttagagctgtagcagcagctacgtagatgtagatctgtagcagagcgctcttgaca 928
Db 3108 GGATGATGTGCTCTACAGGATCTGACATTTATTTGGAAGAGCATTTTGCATA 3049
Qy 929 gggagagagaccagagctgtagaggtgtagcagcagcgtgtagagagctctact 988
Db 3048 GGGAGGAGATGCTAAGGTATAGAGGTGACAGCTTACAGAAAGCAGCGCTGGAGACATATTT 2989
Qy 989 ggcgcgtatgagcgag 1004
Db 2988 GAGAGATGCGGCCAG 2973

RESULT 14
US-08-815-520-1/c
: Sequence 1, Application US/0815520A
: Patent No. 6136569
: GENERAL INFORMATION:
: APPLICANT: Kevin Baker, Austin Gurney
: TITLE OF INVENTION: Method of Selection for Genes Encoding Secreted and
: TITLE OF INVENTION: Transmembrane Proteins
: FILE REFERENCE: P1058
: CURRENT APPLICATION NUMBER: US/08/815,520A
: CURRENT FILING DATE: 1997-02-27
: NUMBER OF SEQ ID NOS: 3
: SEQ ID NO 1
: LENGTH: 7633
: TYPE: DNA
: ORGANISM: Artificial
: FEATURE:
: NAME/KEY: Artificial Sequence
: LOCATION: 1-7633
: OTHER INFORMATION: Amylase yeast expression plasmid
US-08-815-520-1

Query Match 25.8%; Score 356; DB 3; Length 7633;
Best Local Similarity 65.5%; Pred. No. 1.4e-96;
Matches 521; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

Qy 209 catgtcaagaaatcttaccgagaaaggccaaaggacacaaatctgcagctgttagaaa 268
Db 3768 CATGTCAAGAACTATATAGGAAGCTGCTACTACTCTAGCTCTGTTGCTGCCAA 3709
Qy 269 gcttcgcatatgacagagaaagaaacaaatctcgcctccctcgtatgtagcagac 328
Db 3708 GCATTTATATCATCATCAGAAAGCAAAACTTGTGTCTCTCTTGTGATGCTGTC 3649
Qy 329 gtcctagaagctctgtagcagacagcgtgtagcagcagcattgctcgtcgaagac 388

Db 3648 CACCAAGAAATTACTGAGTATGTTGAACCTATTAGTCCCAAAATTTGTTACTAAAC 3589
Qy 389 acatgtcagacatactagcagacttcgacacacagagacagcagcgtcagacac 448
Db 3588 ACATGTGATATCTTGTACATGATTTTTCATGAGGAGCAGATTAAGCCGCTAAAGGCAAT 3529
Qy 449 tgcgctlaagcacaactcagatctcgcagacccagctcgcctcagatctgtagaacac 508
Db 3528 ATCCGCCAAGTACAAATTTTACTCTTCAGACAGAAATTTGCTGACATTTGTAATAC 3469
Qy 509 ggttaagctgacactcctcgcgagctgtagcagacacacagagagtgagacatatac 568
Db 3468 AGTCAAAATTTGACATCTCGCGGTGTATACAGAAATATACAGAAATGCGACACATATAC 3409
Qy 569 tgcacagcgctcaccgagcccggtgtagacgagcgtcgaagagcgtcgaagacac 628
Db 3408 TGCACACGCTGTGGTGGCCAGGATTTGTTAGCGGTTTGAAGCAGGCGGAGAGAAAGT 3349
Qy 629 ctcaagaagaccaggggtgtagatgctgtagcagagcctcctcgaagcctcttgac 688
Db 3348 AACAAAGAACCTAGAGCGCTTTTGTATGTTAGCAGAAATTTGCTATGCAAGGCGCTCCATAC 3289
Qy 689 ggcgagagactataccgagcgctgtagaagctgtagagcctgtagagccttgtag 748
Db 3288 TACTGAGAAATATCTAAGGCTACTGTTGACATTTGCGAAGAGCGCAAAAGATTTGTTAT 3229
Qy 749 cgggttcacgcgacgctgtagatggtggtggtgtagagcgtctgtagagctctgtag 808
Db 3228 CGGCTTTATTTGCTCAAGAGACATGGGTGGAAGAGATGAAAGTTACCATTTGTTAT 3169
Qy 809 gaccccgagggttgccctgtagacagaaagagagcgctgtagcagcagcctacgacgct 868
Db 3168 GACACCGGCTGTGGTGTATGATGACAAAGGAGAGCGATTTGGTCAACAGATATGAACCT 3109
Qy 869 ggttagagctgtagcagcagctacgtagatgtagatctgtagcagagcgctcttgaca 928
Db 3108 GGATGATGTGCTCTACAGGATCTGACATTTATTTGGAAGAGCATTTTGCATA 3049
Qy 929 gggagagagaccagagctgtagaggtgtagcagcagcgtgtagagagcgtctact 988
Db 3048 GGGAGGAGATGCTAAGGTATAGAGGTGACAGCTTACAGAAAGCAGCGCTGGAGACATATTT 2989
Qy 989 ggcgcgtatgagcgag 1004
Db 2988 GAGAGATGCGGCCAG 2973

RESULT 15
US-08-507-455-5
: Sequence 5, Application US/08507455
: Patent No. 5695961
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: BIFUNCTIONAL EXPRESSION SYSTEM
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NIXON & VANDERHAYE, P.C.
: STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
: CITY: ARLINGTON
: STATE: VIRGINIA
: COUNTRY: USA
: ZIP: 22201-4714
COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/507,455
: FILING DATE: 08-SEP-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9303988.1

Wed Apr 4 11:25:31 2001

us-09-582-779a-1.rni

Page 13

/tissue_type="root tips"
 /dev_stage="harvested after 3 days of N-starvation"
 /note="vector: pBluescript PSK; Site_1: EcoRI; Site_2: XhoI; Plants were grown in an aeroponic chamber for 14 days on nitrogen-rich medium followed by 3 days on N-free medium. RNA was extracted from root tips (1-3 cm). cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into uni-zapR vector from StrataGene and packaged using Gigapack Gold packaging extracts. Plasmids containing cDNA inserts were mass-excised from phage stocks using Exsist helper phage and propagated in SOLR cells. Clone ordering and sequencing was performed by the Centre National de Sequencage (Genoscope, Evry, France)."

BASE COUNT 125 a 83 c 118 g 135 t
 ORIGIN

Query Match 7.0%; Score 96.2; DB 28; Length 461;
 Best Local Similarity 54.8%; Pred. No. 4,5e-17;
 Matches 213; Conservative 0; Mismatches 173; Indels 3; Gaps 1;

222 tcttcgacgaagaaggccacagcacaatcgcagctcctagaagaagctcgtgacatg 281
 19 TCATTGGGAGAGGCGCAATCTGTCCAAGATCCACAGCAAGAAAGTTCTCCACATA 78
 282 atgcacgagaagaacacacatctcctcctcctcctcctcctcctcctcctcctc 341
 79 ATCCGCTCAGAGAGAGATATCTGTGTGGCTGCAGATGTGGACATGCGCAATG 138
 342 ctgcagcttagacagacacgctgcagcagacattcgtcgtcgaagacacatcgtcaca 401
 139 GTTGAATATGCGGAGAGAGATGAGACCTGAGATATGCTGCAAGACATCATGTGATTT 198
 402 ctgcagcagctcgcagatcgcagcagcagcagcagcagcagcagcagcagcagcagc 461
 199 TTGCTGATTTTACTCTGATTTTGGCTCTTAAGCTTCTCTCA--TTGCAAGAAACAT 255
 462 aactcagatcctcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 521
 256 AACTCTTAATCTTTCAGAGATCGTAATTTGCTGACATTTGTAACACAGTAACATGCA 315
 522 tactcctcgcagcgtgtacgcgtacgcgtgagcagcagcagcagcagcagcagcagc 581
 316 TATGAGAGAGGATCTTTCACATATTTGATTTGGCTGCATATGTAATGTCACATATTC 375
 582 accgagcagcgtgtgtagtagccggcgctga 610
 376 TCTGGTCTCGAATCGTGAATGACTATA 404

RESULT 13
 AAI66990 391 bp mRNA EST 19-DEC-1996
 LOCUS AAI66990
 DEFINITION zpl2e09.f1 StrataGene fetal retina 937202 Homo sapiens cDNA clone IMAGE:609256 5' similar to gp:J03626 URIDINE 5'-MONOPHOSPHATE SYNTHASE (HUMAN); mRNA sequence.

ACCESSION AAI66990
 VERSION AAI66990.1 GI:1745560
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 391)
 Haller, L., Lennon, G., Becker, M., Bonaldi, M.F., Chiappelli, B., Chissey, S., Dietrich, N., Dubugue, T., Favello, A., Gish, M., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rikkin, L., Rohlfing, T., Schellenger, K., Soares, M.B., Tan, F., Thierly-Meg, J., Trevisan, E., Underwood, K., Wohlmann, P., Watson, R., Wilson, R. and Marra, M.
 Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (9), 807-826 (1996)

MEDLINE 97044478
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estw@wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 361.
 Location/Qualifiers
 1. 391
 /organism="Homo sapiens"
 /db_xref="GDB:4624515"
 /db_xref="taxon:9606"
 /clone="IMAGE:609256"
 /clone_lib="StrataGene fetal retina 937202"
 /sex="mixed"
 /lab_host="SOLR (kanamycin resistant)"
 /note="vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Pooled retinal tissue. Average insert size: 1.0 kb; Uni-zap XR vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGCTTTTCTTTTCTTTT 3'."

BASE COUNT 113 a 77 c 96 g 105 t
 ORIGIN

Query Match 6.9%; Score 95.6; DB 3; Length 391;
 Best Local Similarity 59.6%; Pred. No. 6,4e-17;
 Matches 180; Conservative 0; Mismatches 119; Indels 3; Gaps 1;

255 ccaattgctgaagaagcttcgtcagcagcagcagcagcagcagcagcagcagcagcagc 314
 93 CCAATTGCTGAAGAAGCTTCAGCAGCTTATGCAAAAGAGAGCAATCTGTCTATCT 152
 315 ctgcagctcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 374
 153 GCTGATGTTTTCAGTGGCAGAGAGCTGTGACACTGACAGATGCTTAGGACTAGTATC 212
 375 tctcgtcgaagaacacgtcgcagcagcagcagcagcagcagcagcagcagcagcagcagc 434
 213 TGCATGCTGAAGAGACTGATGATATTTGATGATTTTACTCTGATGATGTA--TGAAG 269
 435 ccgcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 494
 270 GAGTTGATTAACCTGCGAAATGCCATGAGTCTTGTATTTGAAACCGGAAGTTTGA 329
 495 gacattgacacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 554
 330 GATATGAGNAACACAGAGAAAGAGTATGAAAGAGATATCTTTAAATAGCTTCTCTG 389

RESULT 14
 BF044448 570 bp mRNA EST 10-OCT-2000
 LOCUS BF044448
 DEFINITION BP250013A10B4 Soares normalized bovine placenta Bos taurus cDNA clone BP250013A10B4 5', mRNA sequence.

ACCESSION BF044448
 VERSION BF044448.1 GI:10761440
 KEYWORDS EST.
 SOURCE cow.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovine; Bos.
 1 (bases 1 to 570)
 Lewin, H.A., Soares, M.B., Rebeiz, M., Pardini, J., Liu, L. and Larson

Db	600	TGCACACATTATTTCTGTCAGCAAGATGTGATGCTCG	639
RESULT	9		
LOCUS	BE210781		
DEFINITION	BE210781	465 bp mRNA EST	29-JUN-2000
ACCESSION	sos3e04.y1	Gm-cl039 Glycine max cDNA clone	GENOME SYSTEMS CLONE ID
VERSION	Gm-cl039-1687.5'	similar to SW:P4V5.TOBAC Q42942 URIDINE	
KEYWORDS	5'-MONOPHOSPHATE SYNTHASE ;	mRNA sequence.	
SOURCE	BE210781		
ORGANISM	BE210781.1	GI:8827051	
REFERENCE	soybean. Glycine max Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycyne.		
AUTHORS	Shoemaker,R., Kelm,P., Vocklin,L., Expelding,J., Corvelli,V., Rhana, A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wyllie,T., Underwood,K., Steptoe,M., Thelsting,B., Allen,M., Bowers, Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk, R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann, R., Waterston,R. and Wilson.R.		
JOURNAL	Public Soybean EST Project		
COMMENT	Unpublished (1999) Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3322 FAX:(888) 919-3324 or (314) 427-3334 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com High quality sequence stop: 417. Location/Qualifiers 1..465 /organism="Glycine max" /db_xref="taxon:3847" /clone="GENOME SYSTEMS CLONE ID: Gm-cl039-1687" /clone_1lb="Gm-cl039" /issue_type="whole seedling without cotyledons" /lab_host="DH10B" /note="Vector: pluescriptII SK+; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from 2 week old seedlings with the cotyledons removed at the time of harvest. The seedlings for the cultivar Ogden were grown in a growth chamber using germination paper. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the plusscript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker."		
BASE COUNT	124 a	68 c	127 g 146 t
ORIGIN			
Query Match	Best Local Similarity	7.6%; Score 104.8;	DB 105; Length 465;
Matches	215; Conservative	0; Mismatches 162;	Indels 3; Gaps 1;
QY	231	gaaagggccaaggcacacaatcgacgtctgcataaaaagcttcgtcatctgaacgag	290
Db	28	GAGAGGGCCTGAAGCTCTGACAGAATCCAAATGCCAAGAGAGTGTTTGAGATATAGCTGAG	87
QY	231	aagaataaccaatctctgcgtctccctctgatgttgaggacgtctagaagaacttctgagcta	350

Db 88 AAGGAGACTAATCTATGTTTGGCTGCTGATBTTTGGAACCTGCAGCTGAAATTCCTTGAATTT 147

Oy 351 gcaagacacgcctggagaccgcacatctgtctctgtaagaacacatgctgcacabactgaacggac 410

Db 148 GCTGGAAGAGGTTTGGACCTGGAATGTGCTGCTGAAGACATCATGTGATATTTTCCAGAT 207

Oy 411 ttgcacatcctggaagacacagctgaagccgtctgaagaagctctggcgtctaaacacattcatg 470

Db 208 TT---TACTGCTGATTTTTGGCTCTTAACCTTCTCTCGATATGCAGAAAAACATAACTTCTTAA 264

Oy 471 atctctcgagagaccgcaagcttcgcctgcacatctggcaacacggcttgaagctgcagtaactctcc 530

Db 265 ATCTTTGAGGAGATCGTAATATTTCCTGATATTTTGGCAACACACAGTGCACCATGCATATTTGAAGGA 324

Oy 531 ggcctgtacccgtatccgcggagctggggcggatatactcaatgcacacacggcgtccacggccccc 590

Db 325 GGGGTTTTTTCGTAATATTCGATTTGGGCTCATATTAATAATATGCTCACATAAATCTCAGGTCCT 384

Oy 591 ggtgtcatagccggagctgaa 610

Db 385 GGAATTTGATGGATTTAA 404

AA802030	557 bp	mRNA	EST	25-NOV-1998
LOCUS				
DEFINITION	AA802030.Sprlme GM Drosophila melanogaster ovary Bluescript GM037350.Sprlme GM Drosophila melanogaster cDNA clone GM03750 Sprlme similar to L00968: r-1. F89n0003257 PID:g158794 SWISS-PROT:O01637, mRNA sequence.			
ACCESSION	AA802030			
VERSION	AA802030.1	GI:2871149		
KEYWORDS	EST.			
SOURCE	fruit fly.			
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 557) Harvey, D., Hong, L., Evans-Holm, M., Pendleton, J., Su, C., Brokstein P., Lewis, S. and Rubin, G.M. BDGP/HMI Drosophila EST project Unpublished (1997) Contact: Harvey, D. G. M. Rubin-Molecular and Cell Biology University of California Berkeley 539 LSA, Berkeley, CA 94720-3200, USA Fax: 510 643 9947 Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu Plate: 37 row: E column: 2 High quality sequence stop: 395.			
COMMENT				
TITLE				
JOURNAL				
AUTHORS				
REFERENCE				
FEATURES				
source	Location/Qualifiers 1..557 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone="GM03750" /clone_lib="GM Drosophila melanogaster ovary Bluescript" /sex="female" /dev_stage="newly eclosed females: germarium-stage 6" /lab_host="SOLR" /note="Organ: ovary; Vector: Bluescript SK; Site:1: EcorI; Site-2: XhoI; Constructed using Stratagene Zap-cDNA Synthesis kit. Oligo dt-primed and directionally cloned at EcorI and XhoI in Bluescript SK(+/-)"			
BASE COUNT	128 a 152 c 166 g 111 t			
ORIGIN				
Query Match	7.2%	Score 99.4	DB 12	Length 557
Best Local Similarity	58.8%	Pred. No. 5.5e-18		
Matches 191: Conservative	0	Mismatches 111:	Indels 3:	Gaps 1:

427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Seq primer: -40RP from Glibco
High quality sequence stop: 428.

FEATURES

source

1. 586
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl019-6395"
/clone_1bp="Gm-cl019"
/tissue_type="Immature seed coats of greenhouse grown plants"
/lab_host="DH10B (Gibco BRL)"
/note="Vector: pSPORT1 (Life Technologies); Site:1: Not I; Site:2: Sal I; This cDNA library was constructed from mRNA isolated from immature seed coats (200-300 mgs) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a Not I restriction site. Sal I linker adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E. coli. Electromax DH10B host cells (Gibco BRL). This library was constructed by Dr. Lila Vlodin and Dr. Anu Khanna."

BASE COUNT

154 a 101 c 142 g 189 t

ORIGIN

Query Match 7.8%; Score 107.4; DB 94; Length 586;
Best Local Similarity 54.7%; Pred. No. 2.6e-20;
Matches 263; Conservative 0; Mismatches 206; Indels 12; Gaps 2;

QY 280 tgaatcagcagagaagaacaaatctctgcgtcccttgatgtgagcagcttaagaagc 339
DB 12 ttaatgcttgag 71
QY 340 tcttgagagctagacagacagcctgagcagcagacatctgtctgtgagacacagcagc 399
DB 72 tgccttaaaattgcttgagagagagagagagagagagagagagagagagagagag 131
QY 400 tactcagcagactcgacacatcgagacagcagcagcagcagcagcagcagcagcagc 459
DB 132 tttttccagattttactgctgctgctgctgctgctgctgctgctgctgctgctgca 188
QY 460 acacactcctgactcctgagcagcagcagcagcagcagcagcagcagcagcagcagc 519
DB 189 ataaactctttaaattcttgaagagagagagagagagagagagagagagagagagag 248
QY 520 agtactccctcgcgctgtatcgcgagtgagtgagtgagtgagtgagtgagtgagtgag 579
DB 249 aatataag 308
QY 580 tcaacgagcccggtgtgataagcggcgtgaagagagcgtgcaaacctgagcctacaggaac 639
DB 309 tctcagagctcctgag 359
QY 640 ccagagaggtgtgagagcgtgagagagcctctctcagagcctcttgagcagcagagc 699
DB 360 gtaagggctctaatatgcttgctgaatgagcctgtaaccttgagcagagagagag 419
QY 700 atacgcgagcgtcgttgagaaatgagcagagagccttgatcagcagcagcagcagcagc 759
DB 420 atacacacttctgacagtaaaattgctgagagagatcttgcattttgtaattggccttcatct 479
QY 760 c 760
DB 480 c 480

RESULT 8

AM626116

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. 639
/organism="Lycopersicon esculentum"
/cultivar="7496"
/db_xref="taxon:4081"
/clone="cLE218K8"
/clone_1bp="tomato radicle, 5 d post-imbibition, Cornell University"
/tissue_type="radicle"
/dev_stage="seedlings 5 days post-imbibition"
/note="Vector: pBluescript SK(-); Site:1: EcoRI; Site:2: XhoI; Supplier: Tanskley; Tissue supplied by Dave Garvin (USDA-ARS, Ithaca, NY 14850)."
Location/Qualifiers

BASE COUNT

198 a 99 c 166 g 176 t

ORIGIN

Query Match 7.8%; Score 107.2; DB 93; Length 639;
Best Local Similarity 56.0%; Pred. No. 3.1e-20;
Matches 224; Conservative 0; Mismatches 173; Indels 3; Gaps 1;

QY 209 catgtcaacgaatacctaagcagaagaagcagacacacatctgcagcttctgataa 268
DB 243 CAAGACCAAGATTACCAATATGACGAGCGGCAAGACTGCAGAACCTTACTGGAGAGAA 302
QY 269 gctctgcatctgagcag 338
DB 303 GTTGTTGAATCATATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 362
QY 329 gcttagaagcctctgagcagcagcagcagcagcagcagcagcagcagcagcagc 388
DB 363 AGCTGCTGAACCTAGCTGATATGCTGCAACGTTGACCTGAATTCGATGTTAAATAC 422
QY 389 acaatgagacactatgagagagcctcgacatcgagaagcagcagcagcagcagcagc 448
DB 423 ACATGTTGATATATATGCTGATTTTACTCTGACCTTGGTTCTAA---GCTTAGATGAT 479
QY 449 tgcgagcagacacactcagatctctgagagagcagcagcagcagcagcagcagcagcagc 508
DB 480 TGCAAGACAGCAACATCTTAATATTGAGAGATCGTAATATTCCTGACATTTGGAATAC 539
QY 509 ggttaagcgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 568
DB 540 AGTGACATGCAAGCTGAGAGAGAGATCTTCGATTTTGAATTTTGAATTTGAATTTAA 599
QY 569 tgcagagcgtctacgcgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 608

nystatin and cefotaxime in covered crystallization dishes. Roots were harvested. The tissue, total RNA, and poly(A) RNA were prepared, a cDNA library was made in the IV Close lab (Choi, Close, Fenton) at the University of California, Riverside. The cDNA clones were in vivo excised to give pluscript phagemids before normalization and carried out. The mass excision of phagemid library and normalization were done in HT Nguyen lab by D. Zhang at Texas Tech University. Normalization protocol used was that of Soares. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors).

BASE COUNT 144 a 140 c 150 g 147 t

Query Match 9.3%; Score 128.4; DB 108; Length 581;
Best Local Similarity 58.2%; Pred. No. 2e-26;
Matches 269; Conservative 0; Mismatches 181; Indels 12; Gaps 2;

Oy 300 aatctctgctcccttgatgctgagcgtctagaagctcttgagctgacagacag 359
1 AACCTCTGGTGGCCGCGCCGATGTACACAGCAAGAACTCTTGACCTACCAACAG 60
Oy 360 ctggagccgacatctgtctcgaagacacatgacacatctgagcttcgacac 419
61 GTTGGTCCAGATTTGCTGCTGAACCCATGTGATTTCTGATTTTACACCA 120
Oy 420 gagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 479
121 GATTTTGGTGCCAA--GCTCCGTTGATTTGCTGAAGACACACTTTTGTCTTGA 177
Oy 480 gacccgaagctctgagcagcagcagcagcagcagcagcagcagcagcagcag 539
178 GACCCGAATTTGCTGCTGACATGGGGAACGTAATGCAATATGAAGAGGATATTC 237
Oy 540 cgtatcgcgagctggcagcagcagcagcagcagcagcagcagcagcagcagcag 599
238 CCATATACAGCTGGCCGATTTGTCACAGCGCATGTAGTCCGAGCTGATGATCA 297
Oy 600 gccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 659
298 GATGCGTTGAACCTG-----AAGGGTTTGCACAAAGGAAAGGGTTACTTGGT 348
Oy 660 gcaagcctctctctcagcagcagcagcagcagcagcagcagcagcagcagcag 719
349 GCTGAGATAGCGCAGCGACGCAACCTTCTCATGGAGATTACACACGCGACCGTGA 408
Oy 720 atggcgaagcctggcgaagcagcagcagcagcagcagcagcagcagcagcag 761
409 TTGCGCGACCAACATTTGATTTGTTATCGGATTCCTTGTG 450

RESULT 5
BE394063 775 bp mRNA EST 21-JUN-2000
LOCUS 601312432E1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3659383 5',
DEFINITION mRNA sequence.
ACCESSION BE394063
VERSION BE394063.1 GI:9339428
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 775)
AUTHORS NIH-MGC <http://www.ncbi.nlm.nih.gov/MGC/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
COMMENT Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LUCM349 row: p column: 08
High quality sequence stop: 652.
Location/Qualifiers

FEATURES

source

1..775
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3659383"
/clone_lib="NIH_MGC_44"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pORF7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(5). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 197 a 137 c 224 g 197 t

Query Match 8.4%; Score 116; DB 107; Length 775;
Best Local Similarity 56.7%; Pred. No. 8.9e-23;
Matches 279; Conservative 0; Mismatches 200; Indels 13; Gaps 3;

Oy 255 ccaatgctgaagaagctctgagcagcagcagcagcagcagcagcagcagcagcag 314
1 CCATTTGATGAGCCTTCAGCGCTTATGCAAAAGAGACCAATCTGTCTATCT 60
Oy 315 ctgagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 374
61 GCTGATGTTTCACCTGGCCAGAGAGCTG-TGCAGCTACAGATGCTTTAGGACCTAGTATC 119
Oy 375 tctctgctgaagacacatgctgacacatgagcagcagcagcagcagcagcagcagcag 434
120 TGCACTGCTGAAGACTCATGTATTTTGAATGATTTTACTGATGATGCA---TGAAG 176
Oy 435 ccgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 494
177 GAGTTGATTAAGCTTGGCAAAATGCCATGAGTCTTGTATTTGAAGCCCGAATTTGCA 236
Oy 495 gaacttgcaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 554
237 GATATGGAACACAGTGAAGAAACCACTATGAGAGAGATCTTTAAATAGCTTCTG 296
Oy 555 gcgagatctaaccaatgacacagcagcagcagcagcagcagcagcagcagcagcagcag 614
297 GCAGATCTAGTAATAGCTGACAGTGTGCGACAGCTCAGAGATTGGAAGCCCTGCAAG-- 354
Oy 615 gctcggaacatgctcgaagcagcagcagcagcagcagcagcagcagcagcagcagcag 674
355 -----AAGTGGCCCTGCTTTCATCGGGGGGCTCTTATTTAGGAAATGAGCTCC 407
Oy 675 cagggtctcttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 734
408 ACCGCTCCCTGGCCACTGGGACTACACTAGAGACAGCGGTAGAAATGGCTGAGAGACAC 467
Oy 735 gaagactctgt 746
468 TCTGATTTTGTG 479

RESULT 6
BE019969 557 bp mRNA EST 06-JUN-2000
LOCUS BE019969
DEFINITION sm38b04.Y1 Gm-c1028 glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1028-6152 5' similar to SW:PYR5_TOBAC Q42942 URIDINE

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Gossypium hirsutum	Eukaryota: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta	Magnoliophyta, eudicotyledons; core eudicots; Rosidae; eustroids II	Malvales; Malvaceae; Gossypium.	1 (bases 1 to 597)	
Blewitte, M., Metz, E.C., Davy, D.F. and Burr, B.	ESTs from developing cotton fiber	Unpublished (1999)			
Contact: Ben Burr	Biological Department	Brookhaven National Laboratory	Upton, NY 11973, USA	Tel: 516-344-3396	Fax: 516-344-3407
Email: burr@bnl1.bnl.gov					
Seq primer: T3 primer.					
Location/Qualifiers					
1..597					
/organism="Gossypium hirsutum"					
/cultivar="Acala Maxxa"					
/db_xref="taxon:3635"					
/clone_lib="Six-day Cotton fiber"					
/tissue_type="Immature fiber"					
/dev_stage="Six days post anthesis"					
/lab_host="XLI-Blue"					
/note="Vector: pBluescript II KS+"					
BASE COUNT	156 a	117 c	147 g	175 t	2 others
ORIGIN					
Query Match	9.9%	Score 136.4	DB 24	Length 597	
Best Local Similarity	56.2%	Pred. No. 9.6e-29			
Matches	301	Conservative	0	Mismatches 223	Indels 12
					Gaps 2
OY	225	tacgagaaagggccacagccacacaaatcgcagctgtcgtagaagctcttgcatgatg	284		
DB	26	TATCAANAAAGGCGTAAAGTTTCTAAGATCCAACTGGGAAAGCTGTTAGATTAG	85		
OY	285	cacgagaagaaacaatctctgcgctccctctgatlgtgcgagctctagaagctctg	344		
DB	86	GTTAAGAAAGAGATATCTGTTTGGGCGCTGATGTTGCCACTGCTCTGAGTTGCTT	145		
OY	345	gaagctagcagacagccttgagaccgacacatctctgctgtgaagacacatctgcacatctg	404		
DB	146	GATATTTGCGAGAAAGTTGACCTTAATTTGCCATTTGMAAGACCACATGTCACATATTC	205		
OY	405	aaggaatcgcagacatcgaagcagacagtcgaagccgcgcgacgcttgcgctgaagcaac	464		
DB	206	CTTGATTTTACCTCCAGAGATGTCGGTCTTA--GCTTCGGCGGATTTGCACAAATAACCTAAC	262		
OY	465	ttcatgatacttcgagagacccgaagctctgcgtgacatctgcaacagcttaagctgcagtac	524		
DB	263	TTCATGATATTTGAAAGACCGCTAATTTTGGAGACATTCGGTAACACAGTAAATCATATGCACTAT	322		
OY	525	tcctcgcgcgtgtacgatacgcgagatgagcgatataccaatgacacagcgctcacc	584		
DB	333	GAAAGGCGATCTTCCATATATTTGATGATGCGCTGATATGTTAAATGTCACATTAATCTCT	382		
OY	585	ggcccccggtgtgatacgccggtgtgaagagagctgcaaacctgctctcaacagaaaccgag	644		
DB	383	GCTCTTGGAATGTTGATGTTTGAATAATGAA-----GGGTCCTGCTCGTGGTAG	433		
OY	645	gggtctgcagatctgagcaagatctctctctcagagctcttgccgcgcggaacataacc	704		
DB	434	GGGCGTGGCTACTTGCCTGAATAGACTCTGCTGCTGAACCTTGCCACGGGAGGACATACACA	493		
OY	705	gcggagcgtcgttgaatacgcggaagcttgcaagaagacttgtgactgagttcatcgc	760		
DB	494	GCTGCTGCCGTAAATATGCGCAACACACTCCGATTTTGTTCATCGGTTTCATCTC	549		

LOCUS	BE641945	820 bp	mRNA	EST	01-SEP-2000
DEFINITION	Cr12_4_L06_S6	Ceratopteris Spore Library	Ceratopteris richardii		
ACCESSION	CDNA clone Cr12_4_L06 5'	CDNA clone	Ceratopteris richardii		
VERSION	BE641945				
KEYWORDS	BE641945.1	GI:9959615			
SOURCE	EST				
ORGANISM	Ceratopteris richardii				
REFERENCE	Ceratopteris richardii				
AUTHORS	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Filicopsida; Filicopsida; Filicales; Pteridaceae; Ceratopteris				
TITLE	Chatterjee, A., San Miguel, P., Stout, S.C., Banks, J. and Roux, S.J.				
JOURNAL	Expressed sequence tags of cDNA clones from a C. richardii library				
COMMENT	Unpublished (2000)				
	Contact: Roux SJ				
	Section of Molecular Cell and Developmental Biology				
	University of Texas				
	Biology Building, Room 16, Austin, TX 78712, USA				
	Tel: 512 471 4238				
	Fax: 512 232 3402				
	Email: sroux@uts.cc.utexas.edu				
	Plate: Cr12_4 row: L column: 06				
	Seq primer: SP6.				
FEATURES					
Source	Location/Qualifiers				
	1..820				
	/organism="Ceratopteris richardii"				
	/cultivar="Brogan"				
	/db_xref="taxon:49495"				
	/clone="Cr12_4_L06"				
	/clone_lib="Ceratopteris Spore Library"				
	/tissue_type="Gametophyte"				
	/cell_type="Spore"				
	/dev_stage="20 hours after germination initiation"				
	/note="Vector: pcMVSPORT6; EST sequence from cDNA library				
	cdna library constructed from mRNA isolated from C.				
	richardii spores that had developed for 20 hours after				
	their germination had been initiated by white light."				
BASE COUNT	232 a	159 c	219 g	210 t	
ORIGIN					
Query Match	9.5%;	Score 131.2;	DB 110;	Length 820;	
Best Local Similarity	55.6%;	Pred. NO. 3.4e-27;			
Matches 300;	Conservative 0;	Mismatches 28;	Indels 12;	Gaps 2;	
221	atctaacgacagaagaagggccaaagcacacaaatcgcagctgtctagaaagcttcgacat	280			
128	ATCATATATGTGTAAGAGGCAAGCTGTGGCGAGATCCACAGCAAGCGGTGTGGAGAT	187			
281	gatgacgacagaagaacaaatctctgagcctcccttgatgtgcgagctctagaagct	340			
188	AATGAGAGCAAGAAAGAGTAATCTTTCACCTGCTGCAGATGTGATTTGCTGTGACCT	247			
341	tctggagctagcagacagctgaggacgcacattgtctgctgaagacacatgtcgaat	400			
248	CCTTACATATAGCTGTAAGAGGTGGGACCTGAGATTTGTTTGTAAGACACACAGTGCAT	307			
401	actgacggaacttcgacatcgcagacagcagctcaagcgcgtgcgaagccttgcgctaa	460			
308	CTATCTCTATTTCTACTCTCTGATTTTGTGTGCAAA---GCGAGGCTGATTTGCGACAGCA	364			
461	caactlcatlcatlctcgagacgcgaagctcgtctacatcttgcacaacagcttgaagct	520			
365	CACACTTTTGTGATTTTGGAGGATAGGACAGTTTCTGATTTAGGTAACACAGTTACATGCA	424			
521	gtactctcgcgagctgtacgcgtatcgcgagatggcgagatatccaatgcacagcgct	580			
425	GATGAAGGTGGTCTTTTCAAGATTTGTAGAGTGGCGACACATTTCTTAATGCAACATTAT	484			
581	cacgggccccgggtgtgatatagccgggtctgaaggaagccttgcgaatcggcctcaaggaac	640			
485	TCCAGGGCCAGGATATATAGATGAGCTTAAACTCAA-----GGGCTGCGAATATATA	535			

117: em_estp16:**
 118: em_estp17:**
 119: em_estp18:**
 120: em_estp19:**
 121: em_estp20:**
 122: em_estp21:**
 123: em_estp22:**
 124: em_estp23:**
 125: em_estp24:**
 126: em_estp25:**
 127: em_estp26:**
 128: em_estp27:**
 129: em_estp28:**
 130: em_estp29:**
 131: em_estp30:**
 132: em_estp31:**
 133: em_estp32:**
 134: em_estp33:**
 135: em_estp34:**
 136: em_estp35:**
 137: em_estp36:**
 138: em_estp37:**
 139: em_estp38:**
 140: em_estp39:**
 141: em_estp40:**
 142: em_estp41:**
 143: em_estp42:**
 144: em_estp43:**
 145: em_estp44:**
 146: em_estp45:**
 147: em_estp46:**
 148: em_estp47:**
 149: em_estp48:**
 150: em_estp49:**
 151: em_estp50:**
 152: em_estp51:**
 153: em_estp52:**
 154: em_estp53:**
 155: em_estp54:**
 156: em_estp55:**
 157: em_estp56:**
 158: em_estp57:**
 159: em_estp58:**
 160: em_estp59:**
 161: em_estp60:**
 162: em_estp61:**
 163: em_estp62:**
 164: em_estp63:**
 165: em_estp64:**
 166: em_estp65:**
 167: em_estp66:**
 168: em_estp67:**
 169: em_estp68:**
 170: em_estp69:**
 171: em_estp70:**
 172: em_estp71:**
 173: em_estp72:**
 174: em_estp73:**
 175: em_estp74:**
 176: em_estp75:**
 177: em_estp76:**
 178: em_estp77:**
 179: em_estp78:**
 180: em_estp79:**
 181: em_estp80:**
 182: em_estp81:**
 183: em_estp82:**
 184: em_estp83:**
 185: em_estp84:**
 186: em_estp85:**
 187: em_estp86:**
 188: em_estp87:**
 189: em_estp88:**

190: gb_gss25:**
 191: gb_gss26:**
 192: gb_gss27:**
 193: gb_gss28:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	136.4	9.9	597 24	AT731663
2	131.2	9.5	820 110	BE641945
3	128.4	9.3	579 107	BE405348
4	128.4	9.3	581 108	BE446305
5	116	8.4	775 107	BE394063
6	107.6	7.8	557 133	BE019969
7	107.4	7.8	586 94	AW704630
8	107.2	7.8	639 93	AW626116
9	104.8	7.6	465 105	BE210781
10	99.4	7.2	557 12	AA802030
11	97.4	7.1	610 105	BE203967
12	96.2	7.0	461 28	AL369478
13	95.6	6.9	391 3	AA166990
14	95.4	6.9	570 139	BP044448
15	86.4	6.3	987 106	BE300085
16	86.2	6.2	328 31	AV168581
17	85.4	6.2	550 14	AA944506
18	84.2	6.1	420 2	AA128100
19	80.8	5.9	918 109	BE542503
20	79.6	5.8	698 39	AM086516
21	79.2	5.7	549 94	AW755899
22	78.2	5.7	245 40	AM098792
23	75	5.4	649 40	AW116069
24	74.6	5.4	431 2	AA125810
25	74.4	5.4	724 96	AM905992
26	74.2	5.4	668 138	BE951337
27	70.8	5.1	591 89	AM360558
28	67	4.9	247 5	AA328585
29	66.8	4.8	719 26	AI894226
30	65	4.7	148 96	AM851491
31	63.2	4.6	630 8	AA537289
32	60.6	4.4	117 96	AM851496
33	60	4.3	777 140	C97586
34	59.8	4.3	809 135	BE745374
35	59.6	4.3	133 96	AM851587
36	58.8	4.3	617 23	AI665439
37	57.8	4.2	516 93	AM659213
38	57.8	4.2	564 23	AI668476
39	56.8	4.1	279 5	AA295675
40	56.8	4.1	529 22	AI597621
41	55.8	4.0	276 6	AA384977
42	55.2	4.0	636 26	AI862776
43	54.2	3.9	939 108	BE470313
44	54.2	3.9	546 19	AI326762
45	53.8	3.9	498 13	AA919562

ALIGNMENTS

RESULT 1
 LOCUS AT731663 597 bp mRNA EST 11-JUN-1999
 DEFINITION BNLGH110421 Six-day Cotton Fiber Gossypium hirsutum cDNA 5' similar
 to URIDINE 5'-MONOPHOSPHATE SYNTHASE (UMP, mRNA sequence).
 ACCESSION AT731663
 VERSION AT731663.1 GI:5050515
 KEYWORDS EST.
 SOURCE upland cotton.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 3, 2001, 20:11:13 ; Search time 1348.49 Seconds
(without alignments)
7171.231 Million cell updates/sec

Title: US-09-582-779A-1
Perfect score: 1380
Sequence: 1 ctcgagcactcattggaag.....aataatgagtcgttgaattc 1380

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues
Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: em_estbda:*
42: em_estfun:*
43: em_estchum1:*

44: em_esthum2:*
45: em_esthum3:*
46: em_esthum4:*
47: em_esthum5:*
48: em_esthum6:*
49: em_esthum7:*
50: em_esthum8:*
51: em_esthum9:*
52: em_esthum10:*
53: em_esthum11:*
54: em_esthum12:*
55: em_esthum13:*
56: em_esthum14:*
57: em_esthum15:*
58: em_esthum16:*
59: em_esthum17:*
60: em_esthum18:*
61: em_esthum19:*
62: em_esthum20:*
63: em_estln1:*
64: em_estln2:*
65: em_estln3:*
66: em_estln4:*
67: em_estov1:*
68: em_estov2:*
69: em_estp11:*
70: em_estp12:*
71: em_estp13:*
72: em_estp14:*
73: em_estp15:*
74: em_estro1:*
75: em_estro2:*
76: em_estro3:*
77: em_estro4:*
78: em_estro5:*
79: em_estro6:*
80: em_estro7:*
81: em_estro8:*
82: em_estro9:*
83: em_estro10:*
84: em_estro11:*
85: em_estro12:*
86: em_estro13:*
87: gb_est41:*
88: gb_est42:*
89: gb_est43:*
90: gb_est44:*
91: gb_est45:*
92: gb_est46:*
93: gb_est47:*
94: gb_est48:*
95: gb_est49:*
96: gb_est50:*
97: gb_est51:*
98: gb_est52:*
99: gb_est53:*
100: gb_est54:*
101: gb_est55:*
102: gb_est56:*
103: gb_est57:*
104: gb_est67:*
105: gb_est68:*
106: gb_est69:*
107: gb_est70:*
108: gb_est71:*
109: gb_est72:*
110: gb_est73:*
111: gb_est74:*
112: em_estchum21:*
113: em_estchum22:*
114: em_estchum23:*
115: em_estom1:*
116: em_estom2:*

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: Apr11 3, 2001, 21:34:18 ; Search time 2742.82 Seconds
(without alignments)
2574.903 Million cell updates/sec

Title: US-09-582-779a-1

Perfect score: 1380
Sequence: 1 ctcgagcactcattggaag.....aatatgagtcgctgattc 1380

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

GeneBank: *
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_om:*
5: gb_ph:*
6: gb_ph:*
7: gb_ph:*
8: gb_ph:*
9: gb_ph:*
10: gb_ph:*
11: gb_ph:*
12: gb_ph:*
13: gb_ph:*
14: em_fun:*
15: em_fun:*
16: em_fun:*
17: em_fun:*
18: em_fun:*
19: em_fun:*
20: em_fun:*
21: em_fun:*
22: em_fun:*
23: em_fun:*
24: em_fun:*
25: em_fun:*
26: em_fun:*
27: em_fun:*
28: em_fun:*
29: em_fun:*
30: em_fun:*
31: em_fun:*
32: em_fun:*
33: em_fun:*
34: em_fun:*
35: em_fun:*
36: em_fun:*
37: em_fun:*
38: em_fun:*
39: em_fun:*
40: em_fun:*
41: em_fun:*
42: em_fun:*
43: em_fun:*

44: em_hcg2:*
45: em_hcg3:*
46: em_hcg4:*
47: gb_pl3:*
48: gb_pl3:*
49: gb_pl3:*
50: gb_hcg8:*
51: gb_hcg9:*
52: gb_hcg10:*
53: gb_hcg11:*
54: gb_hcg12:*
55: gb_hcg13:*
56: gb_hcg14:*
57: gb_hcg15:*
58: gb_hcg16:*
59: gb_hcg17:*
60: em_hcg4:*
61: em_hcg5:*
62: em_hcg6:*
63: em_hcg7:*
64: em_hcg8:*
65: gb_hcg18:*
66: gb_hcg19:*
67: gb_hcg20:*
68: gb_hcg21:*
69: gb_hcg22:*
70: gb_hcg23:*
71: gb_vil:*
72: gb_vil:*
73: gb_vil:*
74: em_hcg8:*
75: em_hcg9:*
76: em_hcg10:*
77: gb_ph6:*
78: gb_ph7:*
79: gb_ph8:*
80: gb_ph9:*
81: gb_ph10:*
82: gb_ph11:*
83: em_hcg0:*
84: gb_hcg24:*
85: gb_ph8:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380	100.0	1380	81 A94699	A94699 Sequence 1
2	397.4	28.8	1464	6 AB021640	AB021640 Saccharom
3	391	28.3	1021	47 YSACANDIDA	LI361 Candida gla
4	372.8	27.0	1167	47 KLUFA3	Y00454 Kluyveromyc
5	372.8	27.0	1764	47 YSKURA3A	D00431 Yeast Oroti
6	370.4	26.8	2330	81 E11619	E11619 Candida uti
7	370.4	26.8	2330	81 AR065318	AR065318 Sequence
8	369.6	26.8	1469	47 KMURA3GN	221934 K.marxianus
9	369.6	26.8	1469	81 A39345	A39345 Sequence 11
10	365.2	26.5	3750	7 CRU279021	AJ279021 Candida r
11	359.2	26.0	1166	47 YSCODCF	K02207 Yeast (S.ce
12	359.2	26.0	4133	12 AF106619	AF106619 Cloning v
13	359.2	26.0	4342	12 SCU54830	U54830 Synthetic E
14	359.2	26.0	6107	12 SCU54829	U54829 Synthetic E
15	359.2	26.0	6119	12 SCU54828	U54828 Synthetic E
16	359.2	26.0	9344	12 SYNPCS19X	M98046 PCS19 Cloni
17	359.2	26.0	62643	47 SCE871	U18530 Saccharomyc
18	359.2	26.0	115355	84 CEY55D9	AL008876 Caenorhab
19	357.6	25.9	798	7 AF279259	AF279259 Zygosaach
20	357.6	25.9	5225	12 CVU63018	U63018 Cloning vec
21	357	25.9	4303	12 SYNIP352	LI4759 YIP352, yea

QY 1141 gcaacaaatgatacttctgtagatgacatccacacatctgcaagcttgagggtg 1200
|||||
DB 1141 GCAACAAAGTGCATCTTGTTCAGTACATCATCAACACTTGCAGGCTGGGTGG 1200
QY 1201 aagccctcgccggagataatctgfatcccgccatctcggaagtatctgtatgaaa 1260
|||||
DB 1201 AAGCCCTCGCCGGATTAATCTGTAATTCGCCACTTCGTGAATATTGTAATAA 1260
QY 1261 actctcactctgagctagagcacaatactcgacacaaagccctccctacacactcg 1320
|||||
DB 1261 ATCTTCACTTGGGCTAAGTACAGCCATTAAGTACAGCCCTTCTTACACACTTCG 1320
QY 1321 agctggagactaaagtgcacacgaatagcaataatagcaataatgagctgtgaattc 1380
|||||
DB 1321 AGCTGGAGACTAAAGTGCACACGAATACCAATTAATAGCAAAATATGATGCTGAATTC 1380
RESULT 2
AB021640 1464 bp DNA PLN 06-APR-1999
LOCUS Saccharomyces exiguus SeURA4 gene for orotidine-5'-phosphate
DEFINITION decarboxylase, complete cds.
ACCESSION AB021640
VERSION AB021640.1 GI:4589378
KEYWORDS orotidine-5'-phosphate decarboxylase.
SOURCE Saccharomyces exiguus (Strain:Yp74L-3) DNA.
ORGANISM Saccharomyces exiguus
REFERENCE 1 (bases 1 to 1464)
AUTHORS Hisatomi, T., Kodama, T., Toba, H., Moto-oka, M., Watanabe, A. and
Tsudoi, M.
TITLE Molecular cloning and DNA analysis of the orotidine-5'-phosphate
JOURNAL decarboxylase gene from the yeast Saccharomyces exiguus Yp74L-3
REFERENCE 2 (bases 1 to 1464)
AUTHORS Hisatomi, T.
TITLE Direct Submission
JOURNAL Submitted (18-DEC-1998) to the DDBJ/EMBL/Genbank databases. Taisuke
HISATOMI, Fukuyama University, Faculty of Engineering, Department
of Biotechnology; Gakuen-cho 1, Fukuyama, Hiroshima 729-0292, Japan
Tel: +81-849-36-2111 (ex 4627), Fax: +81-849-36-2459)
FEATURES
Source
1..1464
/organism="Saccharomyces exiguus"
/strain="Yp74L-3"
/db_xref="taxon:34358"
488..1294
/gene="SeURA4"
488..1294
/gene="SeURA4"
/product="orotidine-5'-phosphate decarboxylase"
/protein_id="BAA6736.1"
/db_xref="GI:4589378"
/translation="MAFATYERASVHSPVASKLLRLIMASKRTNLCASLDVRYTAE
LSLVQVGYICLTHTVDIDDFSMGKVKPLKLAAYNLLIEDRKFADIGNTVK
LQYSSGVYKIVOMADITNAGVTGSGIYGLKEAEOENSDERGLIMLAESSKSLA
HCEYTGTEIAKTDKOFVFGFIAGRGMGREGEDMILMPGVLIDKGAALGQYR
TVDDVYSGESDIIIVGRGLFGKGRDPTAESRRYRDAGMDAYIKRCNAASN"
BASE COUNT 367 a 337 c 445 g 315 t
ORIGIN
Query Match 28.8%; Score 397.4; DB 6; Length 1464;
Best local Similarity 69.9%; Pred. No. 8.1e-78;
Matches 536; Conservative 0; Mismatches 231; Indels 0; Gaps 0;
QY 223 cttacgcagaagaagcacaacaaatctgcagcttctgtagaagctcttgcatgca 282
|||||
DB 501 CGTACGAGAGAGAGCGGTGGTGCACGCGATCCCTTGCCTGCAAGCTGCGGCGATCA 560

QY 283 tgcacagaagaacaaacaaatctctgcctcccttgatgtgcggaagctctagaagcttc 342
|||||
DB 561 TGCCCTCCCAAGAGACGAACTCTGTGCGCTTGTACGTGCGCACGACTGGCGAGTGC 620
QY 343 tggagctagacagaacacgccttggaacgcacatctgtctgtgaagacacatgacacac 402
|||||
DB 621 TCTCGCTGTGGTACCAAGTGGTCTTACATTTGCTTTTGAAGACTCATGTGGACATCC 680
QY 403 tggcgaactctgacacatcgaggaagcagctcaagcgcctgcgaagctctgagctaacaca 462
|||||
DB 681 TGGACATTTCTCGATGAGGAGGAGCGTGAACCCCTTGAAGAGCTTGTGCGAAGTACA 740
QY 463 actctatctctcagagacgcgaagctgcgtgacacatggaacacgcgttaagctgagct 522
|||||
DB 741 ACTTCTTGAATCTTCGAGACACCGGAATTTGCGGACATCGGGAACACGCTGAAGTTCGACT 800
QY 523 actcctcggcggtgatacgcgtatcgagagtgagcgatatacacaatgacacgcgcgtca 582
|||||
DB 801 ACTCTCTCCGCTGTGTACAGATCGTGCAGTGGGCGACATCAGAAACGCGACGCTGTCA 860
QY 583 ccggcccggtgtgatacgccggcgtggaagagctgcgaacactgcctccacaggaacca 642
|||||
DB 861 CCGGTAGTGGATTTGAAAGGTTGAAAGAGAGCCGCCAGAGAAACAGTGCAGAACCGC 920
QY 643 gggggtgtctgatactgtagcagagctctctcagaagctctcttgccgcgcgagactata 702
|||||
DB 921 GTGGGTTCTGATGCTGTGGGAATTTGCTCGAAGGTTTCACTTGCACACGGGAGTACA 980
QY 703 ccggcgagctgtgataatgagcgaagctgagcgaagacttgatctgagctgatacgcgc 762
|||||
DB 981 CCACGTGGCACTGTAGAGATTTGCAAAACGACACACAGTGTGTTCGGGTTCATTTGCGC 1040
QY 763 agcgtgacatggtgtgagcgtgcagacgcgttgactgactatgaccccgagggtg 822
|||||
DB 1041 AGAGGGGCAATGGGCGGAGACAGAAAGTTTGCAGCTGCTGTATGACGCCGCTGTTG 1100
QY 823 gctctgacgacaaagagagacgcgcctgagcagcaatgacgcgcgtgagatgagctgca 882
|||||
DB 1101 GGCTGAGCATTAAGGCGGATGCTGCTTGACACAGCACTACAGGACCGTGGACGAGTGCCT 1160
QY 883 ggaagagctcgaatgatactgtctgtagaagagcgtctcttgacaaggaagaaccca 942
|||||
DB 1161 CCACGGGCTCGGACATCATCATTTGCGGACAGAGACTTTTGGCAAGGCGAGGCCAAC 1220
QY 943 agctcgaaggtgctccgacacgcgaagcgcgtctggaagctactg 989
|||||
DB 1221 CCGCCGAGAGACGACGCTACAGAGACGCTGGGTGGAGCGCTACTTG 1267
RESULT 3
YSCANDIDA
LOCUS Candida glabrata orotidine-5'-phosphate decarboxylase gene,
DEFINITION complete cds.
ACCESSION L13661
VERSION L13661.1 GI:388364
KEYWORDS orotidine-5'-phosphate decarboxylase.
SOURCE Candida glabrata (strain 85/038) DNA.
ORGANISM Candida glabrata
Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
anamorphic Saccharomycetales; Candida.
REFERENCE 1 (bases 1 to 1021)
AUTHORS Zhou, P., Szczypka, M.S., Young, R.J. and Thiele, D.J.
TITLE A system for gene cloning and manipulation in the yeast Candida
JOURNAL glabrata
MEDLINE Gene 142, 135-140 (1994)
FEATURES
source
1..1021
/organism="Candida glabrata"
/strain="85/038"
/db_xref="taxon:5478"

CDS	85..882	/EC_number="4.1.1.23"
		/codon_start=1
		/product="orotidine-5'-phosphate decarboxylase"
		/protein_id="AA34325.1"
		/db_xref="GI:388365"
		/translation="MSSASVLRQREARSPVASKLKLIMHEKTNLCASIDYTTSEL LKLVDTLGYICLTKTHVDILSDSEBNVTKPKLEMAKRNLIIFEDRFRADIGNTVK LQYTSVYKIAEMADITNAGVTGGIVTGLKQGAETNEPGLMLAEISSGSLA HGEYTKGTVDIAKSDKDFVIGFTIAOKMDGRDGFPMILMTPEVGLDDKDALGQOYR TYDEVFSTGDIILYGRGLFPAKGRDPTEGERYRKAGMDAYLKRIGN"
BASE COUNT	322 a 199 c 216 g 284 t	
ORIGIN		
Query Match	28.38; Score 391; DB 47; Length 1021;	
Best Local Similarity	68.48; Pred. No. 2.2e-76;	
Matches 541; Conservative	0; Mismatches 250; Indels 0; Gaps 0;	
Oy	210 atgtcaacgaatactctaacgagaaagggccaaagcacaatctgcagctgttcagaag 269	
Db	85 ATGTCACAGTGCCTCATATTTACAAAGACCGAGGCTCACCTAGTCCAGTTGCTCCAG 144	
Oy	270 cttctggacatgtagcagagaaacaaatctctgcctcccttgatgtagcagagc 329	
Db	145 TTGCTGAACACTTATCATGAGAGAAACCAATCTTTGCTGCTCTGAGAGTCAACACC 204	
Oy	330 tctagaagcctctgtagcagacagcctgtagccacatctgtctgtagcagagaca 389	
Db	205 ACCAGCGAATTTGTAAGTGTGTTGACACTTTGGCTTCATATTTGCTGCTTAAACACT 264	
Oy	390 catgtcgaatactcgaagacttgacatcgagaagaagtcgaagcagctgacagact 449	
Db	265 CATGTTGATATTTCTATCAGACTCTCTCTCGAAGACCTGTTAAGCCATTAAGAAGATG 324	
Oy	450 gcgagtaagcacaactcaatctctcgagagacgcaagctctgcagacttgacagaca 509	
Db	325 GCAGCCAAACCACTTCTGTTGTTTGAAGTAAATAATTTCCAGACATGTTGTTAACT 384	
Oy	510 gttaaagctgacgtactcctccgagcgtagacgtagcggagtgaggcgagataaccaat 569	
Db	385 GTTAAAGTTGCAATACATCAGTGTCTTACAAATAATTCAGATGGCTGATATAACCAAC 444	
Oy	570 gcaacgagcgtagcagcccggtgtgtagcggagcgtgaagagcgtcgagaactgagcc 629	
Db	445 GCTCATGCTGTACCGGCCAAGGTATCTGCTTAAAGCAAGGTGCTGAAGAAAC 504	
Oy	630 tcaacgaaacagaggggtgtctgtagcgtgcaagactcctctcaagggtcctcttgagc 689	
Db	505 ACCAATGAACCTAGAGGGCTTTTAAATGCTGCAAGTGTCTATTAAGGGCTCTTACT 564	
Oy	690 cgcgagagactaacgcgagggcgtagcgaalgcggaagctgtagcgaagacttgtagc 749	
Db	565 CATGCTGATACACTAAAGGCACCTGATATTTGCTAATCTGATTAAGAGACTTTGTTATT 624	
Oy	750 ggggtatctgcgagcgtagcagtagggtagggcgtagcgaagcgtctgtagctgtagc 809	
Db	625 GGATTCATTGCTCAAAAGATATGGGTGTAGAGACGAGGGCTTTGACTGGCTTAATAATG 684	
Oy	810 accccgggggtctgagcctgtagcagaagaagagcggcgtagcgaagctaacgcaggtg 869	
Db	685 ACTCCGCTGCTGGTGTATGATTAAGCGCATGCACTTGGCTCAACAGTGAAGACTGTT 744	
Oy	870 galgaagctgcgtagcgaagctgtagcagtagcagtagctgtagcagagggccttgtagc 929	
Db	745 GATGAGCTCTTTTCAACCCGCTGACAGATATTTATTCATCTTGGCAGAGGCTATATCCCAAG 804	
Oy	930 gaaagagaccccaaggtctgagaggtacccgtagcgaagcgggtctgtagaggtctactg 989	
Db	805 GGCAGAGATCCAAAGACTGAAGTGAACCTTACAGAAAGCCGCTGGGATGCTTACTTGG 864	
Oy	990 cgcgtagtagg 1000	

Db	865 AAAGATAGC 875	
RESULT 4		
KLURA3	1167 bp	DNA
LOCUS		
DEFINITION	KLURA3 1167 bp DNA	PLN
	Kluveromyces lactis URA3 gene for orotidine-5-phosphate	01-AUG-1996
	decarboxylase.	
ACCESSION	Y00454	
VERSION	Y00454.1	GI:2905
KEYWORDS	decarboxylase; orotidine-5'-phosphate decarboxylase; URA3 gene.	
SOURCE	Kluveromyces lactis.	
ORGANISM	Kluveromyces lactis	
	Fukariyola: Fungi: Ascomycota: Hemiascomycetes: Saccharomycetales;	
	Saccharomycetaceae: Kluveromyces.	
REFERENCE	1 (bases 1 to 1167)	
AUTHORS	Shuster,J.R.	
TITLE	Direct Submission	
JOURNAL	Submitted (24-NOV-1987) Shuster J.R., Chiron Corporation, 4560	
	Horton Street, Emeryville, CA 94608 USA	
REFERENCE	2 (bases 1 to 1167)	
AUTHORS	Shuster,J.R., Moyer,J.D. and Irvine,B.	
TITLE	Sequence of the Kluveromyces lactis URA3 gene	
JOURNAL	Nucleic Acids Res. 15 (20), 8573 (1987)	
MEDLINE	88040483	
COMMENT	*Source: SD 11; library= Plasmid, JSK12.	
FEATURES		
Source	Location/Qualifiers	
	1..1167	
	/organism="Kluveromyces lactis"	
	/db_xref="taxon:28985"	
	178..981	
	/note="decarboxylase (AA 1-267)"	
	/codon_start=1	
	/protein_id="CA68509.1"	
	/db_xref="GI:2906"	
	/db_xref="SWISS-PROT:P07922"	
	/translation="MSTKSYSTRAEHASPVAASKLLRLMDEKTNLCASLDVRSDEL LKLVETLGYICLTKTHVDILDDFSEGVVPLKALAKYKFLFEDRFRADIGNTVK LQYTSVYKIAEMADITNAGVTGGIVTGLKQGAETNEPGLMLAEISSGSLA HGEYTKGTVDIAKSDKDFVIGFTIAONMDGRDGFPMILMTPEVGLDDKDALGQOYR TYDEVYSGSDIILYGRGLFPAKGRDPKVEGERYRNAAGENAVYKRSANH"	
BASE COUNT	347 a 214 c 265 g 341 t	
ORIGIN		
Query Match	27.08; Score 372.8; DB 47; Length 1167;	
Best Local Similarity	67.28; Pred. No. 2.3e-72;	
Matches 527; Conservative	0; Mismatches 257; Indels 0; Gaps 0;	
Oy	203 ttgcacatgtcaacgaatactctaacgagaagggccaagcacaacatctgcagctgagc 262	
Db	171 TTGACCATGTCCCAAAATCATATACCATAGAGTGAACATCAATGCAAGTCCGTTGC 230	
Oy	263 taagaagctcttgagcattgtagcagagaagaacaaacatctctgcctccctgtagt 322	
Db	231 ATCGAAACTTTTACGTTTAAATGATGAAGAAAGAACCAATTTGTGCTTCTTGACGT 290	
Oy	323 gcgagcgtctagaaggtcttgagactagcagaacagcgtgtagcagccacattgtctgct 382	
Db	291 TCGTTGACATGATGAGCTATTTGAACCTTGTGAACCTTGGCTCATATTGCTCTTT 350	
Oy	383 gaagacacatgtagcagactctgacacatgtagcagacagcagctgtagcagcgtgca 442	
Db	351 GAAAAACACAGCTTGATATCTTGGATGATTTACATGATAGAGGTACGTGCTTCATTGAA 410	
Oy	443 gcagctgtagcagcagacacatctgagcctcgagagcgaagctgtagcagctgtagc 502	
Db	411 AGCATTTGGCAGAGAAATATCAAGTTCTTGATATTTGAGGACAGAAATTTCCCATATCGG 470	
Oy	503 caacaggttagagctgtagcagactcctccgagcgtgtagcgtatcgtagcagagtgagat 562	
Db	471 TAAACAGATCAATTTACAATATACATCGGCGCTTTACCGTATGCGAAGATGCTCATAT 530	

ORGANISM Kluyveromyces marxianus
Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
Saccharomycetaceae; Kluyveromyces.
REFERENCE 1 (bases 1 to 1469)
AUTHORS Chapman, J. W., Masters, W., Rouwenhorst, R. J., Toschka, H. Y. and Verbakel, J. M.
TITLE THE USE OF THE KLUYVEROMYCES MARXIANUS INULINASE GENE PROMOTER FOR PROTEIN PRODUCTION
JOURNAL Patent: WO 9413821-A 11 23-JUN-1994;
QUEST INT (NL)
COMMENT Other publication AU 5811194 940704.
FEATURES
source
1. 1469
/organism="Kluyveromyces marxianus"
/strain="CBS 6556"
/db_xref="caxon:4911"
/clone="PKM01"
546. 1349
/gene="URA3"
546. 1349
/gene="URA3"
/standard_name="URA3"
/EC_number="4.1.1.23"
/product="OROTIDINE-5'-PHOSPHATE DECARBOXYLASE"
/protein_id="CAA02438.1"
/db_xref="GI:2295694"
/translation="MSTKSYSERAAHRSPVAKILINMERKSNLCSLDVRYTAE
LKVLEVLPTCLTHVDILDESEFENTIVPLQJAKNHKFLIPEDEKFDIDNTYK
LQYTSGVYRIEMSDITNAHGVYAGIYAKGQAEVETKPRGLIMLAELSSGSLA
HGEYRGVIEIAKSDKDFVIGFIANDNDGEEGDMILMPGVGLDKGALQOQYR
TYDEVVAGSGDIIIVGRGLFAKGRPDVVEGRYRKAQMDATLKRGRSA"

BASE COUNT 443 a 258 c 356 g 412 t
ORIGIN

Query Match 26.8%; Score 369.6; DB 81; Length 1469;
Best local Similarity 66.1%; Pred. No. 1.1e-71;
Matches 534; Conservative 0; Mismatches 274; Indels 0; Gaps 0;

QY 194 accatgacattgcacatctcaagaatctcaagcagaagggccaaagcacacaatc 253
DB 530 AACTAATCAAGAGGAGATGTCGACTGAAGATTACTCGAAGAGCAGCTCATAGAAG 589
QY 254 gccagctgcagaagaatctctgcattgacagagaagaacaacatctctgccttc 313
DB 590 TCCAGTTGCTGCCAACCTTTTAACTTGATGGAAGAGAAAGTCAAACTTATGTCCTTC 649
QY 314 ccttgaatgacagatctcagaagaatctctgagcagcagaacagctggagccgacat 373
DB 650 TCTTGATGTTCTGTAAGAACAGACAGAGTTGTAAGATTAGTTGAGCTTGGCTCCATATAT 709
QY 374 tlgctgcagaagacatgctgcacatctgcagcagcagcagcagcagcagcagcagc 433
DB 710 CTCTCTATTGAAGACACATGATGATCTTGAGAGATTTCACCTTATGAGAAATACCATTTGT 769
QY 434 gccgcgcgcagcagctgcgcgcagcagcagcagcagcagcagcagcagcagcagc 493
DB 770 GCGGTTGAACCAATTGACAGAGAAACAAATTTTGGATTGTAAGACAGAGAAAGTTTGC 829
QY 494 tgcacatgcaacacggttaagctgacgtacccctgcgcgcagcagcagcagcagcagc 553
DB 830 CGCATTTGGGACACCTGTTAAATTAACATACACGCTGCTGTTATACCGTATGCCGAATG 889
QY 554 ggcggaatatcaac 613
DB 890 GTCGTATATCACCAATGACACAGCTGACGTGACGTGACGATTTGTTGCTGTTGAAGA 949
QY 614 ggcctgcgaac 673
DB 950 AGGTGCGGAGGAATTACGAAGAACCTAGAGGGGTTTAATGCTTGGCCGAGTTATCGTC 1009
QY 674 tccagagctcttggcgcgagagacataccgcggcgctgctgaatgacgaagcctgga 733

DB 1010 CAAGGGGCTCTACCGACGCGTGAATACACTCGCGGACCGTGGAATTCACCAAGAGTGA 1069
QY 734 cgaagacttgatgacggtgctcatcgcgcagcagcagcagcagcagcagcagcagcagc 793
DB 1070 TAAAGACTTGTGTATGATTTATGCTCAAAAGATATGCGGAAGAGAAAGGCGCTA 1129
QY 794 tgaactgctacataagccccggggttgctctgagcagcagaagaagagagcctgggcca 853
DB 1130 CGATTGTTGATCATGACGCCAGGCTGTTGCTTTGATGACAAAGGTATGCTTTGGGACA 1189
QY 854 gcaatcacacacagctgagatgagctgcacagcagcagcagcagcagcagcagcagcagc 913
DB 1190 ACAATPACAGAACTGTGATGAAGTTGTCGCGGTGATCAGACATCATTTATGTTGTTAG 1249
QY 914 aggcctcttgacaagaggaagagagagagagagagagagagagagagagagagagag 973
DB 1250 AGGCTTTTCGCAAGGAGAGATCCTGTATGAGAGAGTGAAGATACGAAAGGCGGG 1309
QY 974 ttggagagcctacttgccgcagcagcagcagcagcagcagcagcagcagcagcagc 1001
DB 1310 ATGGAGCGCTTACTTGAAGAGATAGGC 1337

RESULT 10
CRU279021 3750 bp DNA PLN 16-SEP-2000
LOCUS CRU279021
DEFINITION Candida rugosa, putative URA3 gene for putative
orotidine-5'-phosphate decarboxylase and partial ACP gene for acyl
carrier protein.
ACCESSION AJ279021 GI:10185021
VERSION AJ279021.1 GI:10185021
KEYWORDS acp gene; acyl carrier protein; orotidine-5'-phosphate
decarboxylase; URA3 gene.
SOURCE
ORGANISM
Candida rugosa.
Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
anamorphic Saccharomycetales; Candida.
REFERENCE 1 (bases 1 to 3750)
AUTHORS Biasio, W.
JOURNAL Thesis (2000) University of Vienna, Department of Biochemistry and
Molecular Cell Biology
2 (bases 1 to 3750)
REFERENCE
AUTHORS Biasio, W.
TITLE Direct Submission
JOURNAL Submitted (11-SEP-2000) Biasio W., Department of Biochemistry and
Molecular Cell Biology, University of Vienna, Dr Bohrergasse 9,
A-1030 Vienna, Austria
location/Qualifiers
1. 3750
/organism="Candida rugosa"
/db_xref="taxon:5481"
1. 37
/gene="ACP"
1. 37
/gene="ACP"
/gene="ACP"
/codon_start=2
/transl_table=12
/product="acyl carrier protein"
/protein_id="CAC08812.1"
/db_xref="GI:10185023"
/translation="VDYVLRQPDVAV"
1123. 1920
/gene="putative URA3"
1123. 1920
/gene="putative URA3"
/EC_number="4.1.1.23"
/codon_start=1
/transl_table=12
/codon="seq",aa:Ser)
/product="putative Orotidine-5'-phosphate decarboxylase"
/protein_id="CAC08811.1"
/db_xref="GI:10185022"

repeat_unit	1..38	/rpt_type=terminal
misc_recomb	93..126	/note="loxR site"
gene	135..851	/organism="synthetic construct"
CDS	135..851	/gene="GFP"
	135..851	/gene="GFP"
	135..851	/codon_start=1
	135..851	/product="green fluorescent protein"
	135..851	/protein_id="AAB01173.1"
	135..851	/db_xref="GI:1335938"
	135..851	/translation="MSKKEELFTGVPLIVELDDGVNGHKFSVSGEGEGDATYGLKTLITL
	135..851	KEFTGKGLPVPWPLIVTFFSYGVOCFSRPYDHHKOHFFESAMPEGVORTLFFPKIK
	135..851	IGNKTRAEVKEEGEDTLVNRTELKGDIPKEGNGNLGHLNINSHNYTIDAKQNG
	135..851	IKVNPKRINIEDSGVOLADHYQQNTPIGDSVILLPNHYLTQSAISKDPERRHH
	135..851	VLVEVPIAGITTHCMDELK"
gene	1688..2491	/gene="URA3"
CDS	1688..2491	/gene="URA3"
	1688..2491	/note="ura3p from S. cerevisiae"
	1688..2491	/codon_start=1
	1688..2491	/transl_table=1
	1688..2491	/product="Omp decarboxylase"
	1688..2491	/protein_id="AAB01174.1"
	1688..2491	/db_xref="GI:1335939"
	1688..2491	/translation="MSKATYKERAATHPSPPAAKLFNIMHKONTLCASLDVPTTEL
	1688..2491	LELVAGYAPKICLKLTHTVDITLTDPSMGCTVYPLALSAKYNFLFEDRKADIGNTVK
	1688..2491	LOYAGYARIEMADITRNAHGVSPGIVSGIKQAAEYTKPRGLMLAEISGSLTA
	1688..2491	TGETTKGTVDIAKSDKDPVIGFIAQRMGSGDEGYDWLIMTPGVGLDDKGDALQOQYR
	1688..2491	TVDVVSTGSDIITVGRGLFAKGRDAVEGERYRKAGWEAYLRRCGOON"
gene	2684..3874	/gene="tet"
CDS	2684..3874	/gene="tet"
	2684..3874	/note="tet"
	2684..3874	/codon_start=1
	2684..3874	/transl_table=1
	2684..3874	/product="tetracycline efflux protein"
	2684..3874	/protein_id="AAB01195.1"
	2684..3874	/db_xref="GI:1335940"
	2684..3874	/translation="MKSNNALIVITLGTVDVAVSIGLVMPLPGLLRDIYHSDSIASH
	2684..3874	YGVLLALYALMOPICAPVIGALSDPRRPPVLLASLGLATIDYIMATPTVWLITLH
	2684..3874	RIVAGITGATGAVAGATVADITDDEDARHFGILMSACFGVGVAGVPAAGLLGISLH
	2684..3874	APFLAALVNLNLILGCFLMQESHKGERRMPLRAPNPVSSFMWAGMTIVAALEML
	2684..3874	FETVLGVGPAAALVIFGEDRFWSATMIGLSLAVGILHALAQAVTPTARFEE
	2684..3874	KQALIAAGMADALGVYPLAFTRGMAFPIMILLASGIGMPLAQMILSPOVDDHGG
	2684..3874	QLOSLAALTSITSTGTPLVTAIYASASATWNGLAMIYGAALTYLGPLLRGAMSR
	2684..3874	ATSS"
misc_recomb	4101..4126	/complement(4101..4126)
	4101..4126	/note="n3 res site"
misc_recomb	4161..4194	/organism="synthetic construct"
	4161..4194	/note="loxP site"
CDS	4161..4194	/organism="synthetic construct"
	4161..4194	/note="from influenza virus"
	4161..4194	/codon_start=1
	4161..4194	/transl_table=1
	4161..4194	/product="hemagglutinin triple-epitope"
	4161..4194	/protein_id="AAB01176.1"
	4161..4194	/db_xref="GI:1335941"
	4161..4194	/translation="YPRDVPDYAGYPRDVPDYAGSYPRDVPDYA"
repeat_unit	4305..4321	/rpt_type=terminal
BASE COUNT	1153 a 1014 c 1084 g 1091 t	

Best Local Similarity - 65.7%; Pred. No. 2e-69; Indels 0; Gaps 0
Matches 523; Conservative 0; Mismatches 273;

LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE
209	catgccaacgaacattcttaacgaagaagggcccaagagcacaacattctgcaggtctcttaaga	268				
1687	CATGTCGGAACACTACATATTAAGGAGACGTGCTGCTACTACTACATCTTACGTCCTGTGGCCAA	1746				
269	gctctcgtgacattgbbgcaggaagaacaaactctcgtctccctctgaatgctcgagc	328				
1747	GCTATTTAAATTCATATGCAACGAAAGCAAAACAACCTGTGTGCTTCAATTTGATGTCTGAC	1806				
329	gtctcgaagaagctcttcggagctagaagacacagcctgggaacgcacattgctcgtctgaagc	388				
1807	CACCAAGAAATTACTGCGATTAGTTGAGACATTATAGCTCCAAATTTGTTTACTTAANAAC	1866				
389	acatgtgcacatactgcagcagcacttcgacatctgcagaacagacatcaaacgcctgcagcagct	448				
1867	ACATGTGCAATTCCTTTCGACTGATTTTTCATGAGAGGCGACAGTTAAACCGCTTAAAGCAATT	1926				
449	tcggcgtcgaagacacacttcaatgactctcggagacgcgcaatgctgcgtatattggcaaac	508				
1927	ATCCGCCAAGATACATATTTTCTTACTCTTCGAGACAGCAAAATTTGCTGCACATTTGGTAATAC	1986				
509	ggctaaagcttcgaagctcctcctccgcgctgtaaccgtatccgcggagtcggcggaatattacca	568				
1987	AGTCAAAATTTGCAGTACTCTGGGGGTATACAGAAATAGCAGATGGCGACATTTACGAA	2046				
569	tgcaacagcggtaccgcggcccggtctgtataagccgggctcgaaggaagctcgcgaaacctgc	628				
2047	TGCACACGGTGTGGTGGCCAGGTATTTGACGGCTTTGAAAGCAGCGCGCGAAGAAAGT	2106				
629	ctcacaggaacccagggaggtctgcgtatgctctgcagagcctctctcgaaggtctcttgcc	688				
2107	AACAAAGGAACCTTAGAGGCGCTTTTGATGTTAGCAGAAATTTGTATGCAAGCGCTCCCTAGC	2166				
689	gcggggagacatactccgcgggctgcgtctgtaaatggcggaagctcggagcaagactttgcat	748				
2167	TACTGGAAATATATCTTAAAGGTACTGTTGACATTTGCGAAGAGGAGAAAGATTTTGTAT	2226				
749	cgaggctcacgcgagcagctgacatgagctgggcgtgcagacggcctcttgacttgctcatcat	808				
2227	CGGGTTTATTCCTCAAGAGACATGGGTGCAGATGAAGGTTACGATTTGTTGATTTAT	2286				
809	gaaccccgagggtcttgctctgtaacgacaagaagagagcgccctgggcccagcagttacgcagct	868				
2287	GACACCCGGGTGGGTTTAAATGACAAAGGAGACGCGATTGGGTCAACATATTAAGAACCGT	2346				
869	ggaatgaggtctgcagcagctacacgatgacatctgttcgcaagagggcctctcttgacaa	928				
2347	GGATGATGTGTCTCTACAGAGATCTGACATTTATTTGTTGCAAGAGGACTTTTGGCAAA	2406				
929	gggaagaagaccccaaggtctcgaagggtcccgctacgcgcaagcggtctgggaagcttact	988				
2407	GGGAAGGAGATGCTTAAAGGTACAGGGTGAACGTTTACAGAAAGCAGCGCTGGGAAGCATATTT	2466				
989	gcgcgcgatacggcgag	1004				
2467	GAGAAAGATGGCGCCAG	2482				
RESULT 14						
LOCUS	SCU54829					
DEFINITION	SCU54829 6107 bp DNA SYN 21-DEC-1999					
ACCESSION	U54829					
VERSION	U54829.1					
KEYWORDS	synthetic construct.					
SOURCE	synthetic construct.					
ORGANISM	artificial sequence.					
REFERENCE	1 (bases 1 to 6107)					


```
Db 3524 GCTATTATATTCATGCAAGAAACAAACTTGCTGCTTCAATGATGCTTAC 3583
OY 329 gctagaagctctcgtgagctagacagacgcttggaacgcacattgctcgtgaagac 388
Db 3584 CACCAAGGAATTTACTGAGGTAGTTGAAAGCATTTAGTCCCAAAATTTGTTACTAAAC 3643
OY 389 acatgctgacatcagcagcctcgcacatcgcagcagcagtcgaagccgctgcagcagct 448
Db 3644 ACATGCTGATATCTTGAATGATTTTCCATGAGGGCCACAGTTAAGCCGCTAAAGGCATT 3703
OY 449 tgcgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 508
Db 3704 ATCCGCCAAGTACAAATTTTCTTACTTCTGAAAGACAGAAATTTGCTGACATTTGTAATAC 3763
OY 509 ggtlaagctgacatcctcctcgtgacgctgacgctgacgctgacgctgacgctgacgct 568
Db 3764 ACTCAATTTGCAATCTGCGGGGTATACAGAAATAGCAAAATGCGGACATTTACGAA 3823
OY 569 tgcacagcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 628
Db 3824 TGCACACGCTGCTGAGGCGCCAGGATTTGTTAGCGCTTGAAGCAAGCGCGGGAAGAGT 3883
OY 629 ctcacaggaacccagggggtgctgctgctgctgctgctgctgctgctgctgctgctgct 688
Db 3884 AACCAAGGAACTTACAGGCTTTTATGATGTAAGCAATTTGATGCAAGGCTCCCTAGC 3943
OY 689 gcgcgagagactacacgagcgcgtcgttgaatlgcgaagctgacgagacttgcgt 748
Db 3944 TACTGAGAAATATCTAAGGGTACTGTTGACATTCGGAAGCGCAAAAGATTTTGTAT 4003
OY 749 cgggttcacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 808
Db 4004 CGGCTTTATTTGCTCAAAAGACATGCTGGAAGAGATGAAGTTAGATTGCTTAT 4063
OY 809 gaccccgagggttggccttggacgacagagagagcgcgcgcgcgcgcgcgcgcgcgcgc 868
Db 4064 GACACCGGCTGGGTTAGATGACAGGAGAGCAGTTGGGTCAACAGTATAGAACCGT 4123
OY 869 ggaatgaggtcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 928
Db 4124 GGATGATGCTGCTTACAGGATCTGACATTTATTTGTTGAGAGAGAGCTATTTGCANA 4183
OY 929 ggaagagagacccagagctcgcagcagcagcagcagcagcagcagcagcagcagcagc 988
Db 4184 GGAAGAGATGCTAGAGTGAAGGAGGAGCTTACAGAAAGCAGGCTGGAAAGCATATTT 4243
OY 989 gcgcgctatlggcgag 1004
Db 4244 GAGNAGATCGGCCAG 4259
```

Search completed: April 4, 2001, 00:00:08
Job time: 8750 sec